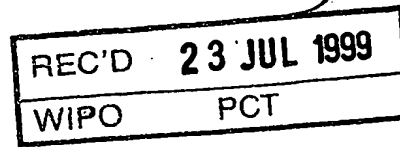


EP 99/3889



Bescheinigung

09/701586

Die BASF Aktiengesellschaft in Ludwigshafen/Deutschland hat eine Patentanmeldung unter der Bezeichnung

"Neue Poly(ADP-ribose)polymerase-Gene"

am 5. Juni 1998 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig die Symbole C 12 N, C 07 H und A 01 K der Internationalen Patentklassifikation erhalten.

München, den 18. Juni 1999

Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Aktenzeichen: 198 25 213.7

Ebert

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

M

200599

Neue Poly(ADP-ribose)polymerase-Gene

Beschreibung

5

Die vorliegende Erfindung betrifft neue Poly(ADP-ribose)polymerase (PARP) Gene und die von ihnen abgeleiteten Proteine; Antikörper mit Spezifität für die neuen Proteine; pharmazeutische und gentherapeutische Mittel, welche erfindungsgemäße Produkte enthalten; Verfahren zur analytischen Bestimmung der erfindungsgemäßen Proteine und Nukleinsäuren; Verfahren zur Identifizierung von Effektoren oder Bindungspartnern der erfindungsgemäßen Proteine und Verfahren zur Bestimmung der Wirksamkeit solcher Effektoren.

15

Die primäre, physiologische Funktion von PARP (EC 2.4.2.30) (teilweise auch bezeichnet als PARS, Poly(adenosin-5'-diphosphoribose)synthetase) scheint in deren Beteiligung an einem komplexen Reparaturmechanismus zu bestehen, den Zellen zur Reparatur von DNA-Strangbrüchen entwickelt haben. Die primäre zelluläre Reaktion auf einen DNA-Strangbruch scheint dabei in der PARP-katalysierten Synthese von Poly(ADP-ribose) aus NAD^+ zu bestehen (vgl. De Murcia, G. et al. (1994) TIBS, 19, 172).

PARP besitzt eine modular aufgebaute Molekülstruktur. Drei funktionale Hauptelemente wurden bisher identifiziert: eine N-terminale 46kDa große DNA-Bindungsdomäne; eine zentrale 22kDa Automodifikationsdomäne, an welche Poly(ADP-ribose) angehängt wird, wobei die DNA-Affinität mit zunehmender Elongation abnimmt; und eine C-terminale 54 kDa große NAD^+ -Bindungsdomäne. Lediglich in PARP aus *Drosophila* wurde eine Leucin-Zipper-Region innerhalb der Automodifikationsdomäne festgestellt, welche auf mögliche Protein-Protein-Interaktionen hinweist. Alle bisher bekannten PARPs sind nur als Homodimere aktiv.

35

Der hohe Organisationsgrad des Moleküls spiegelt sich wieder in der starken Konservierung der Aminosäuresequenz. So wurde für PARP aus Mensch, Maus, Rind und Huhn eine 62%-ige Konservierung der Aminosäuresequenz festgestellt. Größere strukturelle Unterschiede bestehen zu PARP aus *Drosophila*. Die einzelnen Domänen selbst weisen wiederum Cluster mit erhöhter Konservierung auf. So enthält die DNA-Bindungsregion zwei sogenannte Zink-Finger als Unterdomänen (umfassend Motive des Typs $\text{CX}_2\text{CX}_{28/30}\text{HX}_2\text{C}$), die an der Zn^{2+} -abhängigen Erkennung von Strangbrüchen beteiligt sind. Die C-terminale katalytische Domäne umfaßt einen Block von etwa 50 Aminosäuren (Reste 859-908), der unter Vertebraten zu 100% konserviert ist. Dieser Block bindet das natürliche Substrat NAD^+ und

H

200899

2

bedingt somit die Synthese von Poly(ADP-ribose) (vgl. de Murcia, a.a.O.). Insbesondere das GX₃GKG-Motiv ist für PARP in diesem Block charakteristisch.

- 5 Der oben beschriebenen positiven Funktion steht eine pathologische in zahlreichen Krankheiten (Schlaganfall, Herzinfarkt, Sepsis etc.) gegenüber. PARP ist in den Zelltod infolge von Ischemien des Gehirns (Choi, D.W., (1997) Nature medicine, 3, 10, 1073), des Myokards (Zingarelli, B., et al (1997), Cardiovascular Research, 36, 205) und auch des Auges (Lam, T.T. (1997), Res. Comm. in Molecular Pathology and Pharmacology, 95, 3, 241) involviert. Auch bei septischem Schock wurde eine durch Entzündungs-Mediatoren ausgelöste PARP-Aktivierung beobachtet (Szabo, C., et al. (1997), Journal of Clinical Investigation, 100, 3, 723). Da-
15 bei geht eine Aktivierung von PARP mit einem starken Verbrauch an NAD⁺ einher. Da für die Biosynthese von einem Mol NAD⁺ vier Mole ATP verbraucht werden, nimmt die zelluläre Energieversorgung drastisch ab. Zelltod ist die Folge.
- 20 Als PARP-Inhibitoren werden in der oben genannten Fachliteratur Nicotinamid und 3-Aminobenzamid beschrieben. 3,4-Dihydro-5[4-1(1-piperidinyl)butoxy]-1(2H)-isochinolon ist aus Takahashi, K., et al (1997), Journal of Cerebral Blood Flow and Metabolism 17, 1137 bekannt. Weitere Inhibitoren werden z.B. be-
25 schrieben in Banasik, M., et al. (1992) J. Biol. Chem., 267, 3, 1569 und Griffin, R.J., et al. (1995), Anti-Cancer Drug Design, 10, 507.

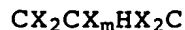
- Als hochmolekulare Bindungspartner für humanes PARP ist unter anderem das Base Excision Repair (BER) Protein XRCC1 (X-ray repair cross-complementing 1) beschrieben worden, das über ein Zink-Finger-Motiv und ein BRCT (BRCA1 C-Terminus) Modul (Aminosäuren 372-524) bindet (Masson, M., et al., (1998) Molecular and Cellular Biology, 18,6, 3563).

- 35 Aufgrund der vielfältigen physiologischen und pathologischen Funktionen von PARP stellte sich die Aufgabe, neue PARP-Homologe bereitzustellen. Die Bereitstellung von homologen PARP's wäre nämlich von besonderer Bedeutung für die Entwicklung neuer Wirkstofftargets bzw. neuer Wirkstoffe, um Diagnose und/oder Therapie von Krankheitszuständen zu verbessern, in welche PARP, PARP-Homologe oder davon abgeleitete Substanzen involviert sind.
40

- Diese Aufgabe wurde überraschenderweise gelöst durch Bereitstellung von PARP-Homologen, gekennzeichnet durch eine Aminosäuresequenz, die
45 a) eine funktionale NAD⁺-Bindungsdomäne

und

- b) insbesondere im N-terminalen Sequenzbereich, d.h. im Bereich der ersten 200, wie z.B. im Bereich der ersten 100, N-terminalen Aminosäuren, keine PARP-Zink-Finger-Sequenzmotive der allgemeinen Formel



aufweist, worin

- m für einen ganzzahligen Wert von 28 oder 30 steht und die Reste X unabhängig voneinander für eine beliebige Aminosäure stehen;
und die funktionalen Äquivalente davon.

- Wesentliches Kennzeichen der erfindungsgemäßen PARPs ist somit das Vorhandensein einer funktionalen NAD⁺-Bindungsdomäne, welche im C-terminalen Bereich der Aminosäuresequenz (d.h. etwa im Bereich der letzten 400, wie z.B. der letzten 350 oder 300) C-terminalen Aminosäuren) liegt, in Kombination mit einer keine Zink-Finger-Motive aufweisenden N-terminalen Sequenz. Da die Zink-Finger-Motive in bekannten PARPs vermutlich zur Erkennung der DNA-Bruchstellen beitragen, ist anzunehmen, daß die erfindungsgemäßen Proteine, wenn überhaupt, mit DNA in anderer Weise wechselwirken.

- Die funktionale NAD⁺-Bindungsdomäne (d.h. katalytische Domäne) bindet das Substrat für die Poly-ADP-Ribose-Synthese. In Übereinstimmung mit bekannten PARPs ist insbesondere das Sequenzmotiv GX₁X₂X₃GKG, worin G für Glycin steht, K für Lysin steht und X₁, X₂ und X₃ unabhängig voneinander für eine beliebige Aminosäure stehen, zu finden. Wie jedoch ein Vergleich der Aminosäuresequenzen der NAD⁺-Bindungsdomänen erfindungsgemäßer PARP-Moleküle mit bisher bekanntem humanem PARP (im folgenden bezeichnet "humanes PARP1") überraschenderweise zeigt, weichen die erfindungsgemäßen Sequenzen von der bekannten Sequenz für die NAD⁺-Bindungsdomäne deutlich ab.

- Einer erfindungsgemäß bevorzugten Gruppe von PARP-Molekülen ist folgendes allgemeines Sequenzmotiv in der katalytischen Domäne vorzugsweise gemeinsam:

- PX_n(S/T)GX₃GKGIYFA, insbesondere
(S/T)XGLRIPX_n(S/T)GX₃GKGIYFA, vorzugsweise
LLWHG(S/T)X₇IL(S/T)XGLRIPX_n(S/T)GX₃GKGIYFAX₃SKSAXY

- worin (S/T) die alternative Besetzung dieser Sequenzposition mit S oder T beschreibt und n für einen ganzzahligen Wert von 1 bis 5 steht und die Reste X unabhängig voneinander für eine beliebige

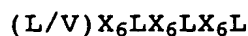
M 29.05.99

4

Aminosäure stehen. Letztes Motiv wird auch als "PARP-Signature" Motiv bezeichnet.

Die Automodifikationsdomäne ist in den erfindungsgemäßen PARPs vorzugsweise ebenfalls ausgebildet. Sie kann etwa im Bereich von etwa 100 bis 200 Aminosäuren vor dem N-terminalen Ende der NAD⁺-Bindungsdomäne liegen.

Eine Gruppe bevorzugter erfindungsgemäßer PARP-Homologer ist außerdem dadurch gekennzeichnet, daß sie N-terminal zur NAD⁺-Bindungsdomäne (d.h. etwa 30 bis etwa 80 Aminosäuren näher am N-Terminus) ein Leucin-Zipper-artiges Sequenzmotiv der allgemeinen Formel



- umfaßt, worin (L/V) für die alternative Besetzung dieser Sequenzposition mit L oder V steht und die Reste X unabhängig voneinander für eine beliebige Aminosäure stehen. Die erfindungsgemäß beobachteten Leucin-Zipper-Motive unterscheiden sich hinsichtlich ihrer Lage deutlich von den für PARP aus Drosophila beschriebenen. Leucin-Zipper können zu Homodimeren (zwei PARP-Moleküle) oder Heterodimeren (ein PARP-Molekül mit einem davon verschiedenen Bindungspartner) führen.
- Die erfindungsgemäßen PARP-Homologen umfassen vorzugsweise außerdem N-terminal zum oben genannten Leucin-Zipper-artigen Sequenzmotiv, d.h. etwa 10 bis 250 Aminosäurereste näher am N-Terminus, wenigstens ein weiteres der folgenden Teilsequenz-Motive:

- | | | |
|----|---|---------------|
| 30 | LX ₉ NX ₂ YX ₂ QLXDX _b WGRVG, | (Motiv 1) |
| | AX ₃ FXKX ₄ KTXNXWX ₅ FX ₃ PXK, | (Motiv 2) |
| | QXLIX ₂ IX ₉ MX ₁₀ PLGKLX ₃ QIX ₆ L, | (Motiv 3) |
| | FYTXIPHXFGX ₃ PP, | (Motiv 4) und |
| | KX ₃ LX ₂ LXDIEXAX ₂ L | (Motiv 5), |

- 35 worin b für den ganzzahligen Wert 10 oder 11 steht und die Reste X unabhängig voneinander für eine beliebige Aminosäure stehen. Am meisten bevorzugt sind diese Motive 1 bis 5 in der genannten Reihenfolge gleichzeitig vorhanden, wobei Motiv 1 dem N-Terminus am nächsten liegt.

Auf das oben genannte PARP-Signature Motiv folgt in den erfindungsgemäßen Proteinen wenigstens ein weiteres der folgenden Motive:

- | | | |
|----|---|---------------|
| 45 | GX ₃ LXEVALG | (Motiv 6) |
| | GX ₂ SX ₄ GX ₃ PX _a LXGX ₂ V | (Motiv 7) und |
| | EYX ₂ YX ₃ QX ₄ YLL | (Motiv 8) |

H

2005.99

5

worin a gleich 7 bis 9 und X jeweils für einen beliebige Aminosäure steht. Am bevorzugtesten sind die drei C-terminalen Motive gleichzeitig und in der genannten Reihenfolge ausgebildet, wobei Motiv 8 dem C-Terminus am nächsten liegt.

5

Ein bevorzugter Aufbau einer erfindungsgemäßen PARP-Struktur kann schematisch wie folgt beschrieben werden:

Motive 1 bis 5/Leucin-Zipper/PARP-Signature/Motive 6 bis 8

10

wobei zwischen den Einzelmotiven weiterer Aminosäurereste, wie z.B. bis zu 40, und am N-Terminus und/oder am C-Terminus weitere Aminosäurereste, wie z.B. bis zu 80, angeordnet sein können.

- 15 Erfindungsgemäß besonders bevorzugte PARP-Homologe sind die Proteine humanPARP2 und humanPARP3 und die funktionalen Äquivalente davon. Das als humanPARP2 bezeichnete Proteine umfaßt 570 Aminosäuren (vgl. SEQ ID NO:2). Das als humanPARP3 bezeichnete Protein existiert möglicherweise in zwei Formen. Typ 1 umfaßt dabei 533
- 20 Aminosäuren (SEQ ID NO:4) und Typ 2 umfaßt 540 Aminosäuren (SEQ ID NO:6).

Ein weiterer Gegenstand der Erfindung betrifft Bindungspartner für die erfindungsgemäßen PARP-Homologen. Diese Bindungspartner

- 25 sind vorzugsweise ausgewählt unter

- a) Antikörpern und Fragmenten, wie z.B. Fv, Fab, (Fab)'₂, davon
- b) proteinartigen Verbindungen, welche, z.B. über die obige Leucin-Zipper-Region oder einen anderen Sequenzabschnitt, mit PARP wechselwirken, und
- 30 c) niedermolekularen Effektoren, welche eine biologische PARP-Funktion, wie z.B. die katalytische PARP-Aktivität, d.h. die NAD⁺-verbrauchende ADP-Ribosylierung, oder die Bindung an ein Aktivatorprotein oder an DNA modulieren.

- 35 Ein weiterer Gegenstand der Erfindung betrifft Nukleinsäuren, umfassend

- a) eine, für wenigstens ein erfindungsgemäßes PARP-Homologes kodierende, Nukleotidsequenz oder die komplementäre Nukleotidsequenz davon;
- 40 b) eine Nukleotidsequenz, die, vorzugsweise unter stringenten Bedingungen mit einer Sequenz gemäß a) hybridisiert; oder
- c) Nukleotidsequenzen, die durch Entartung des genetischen Codes von den in a) und b) definierten Nukleotidsequenzen abgeleitet sind.

45

M

25.05.99

6

Insbesondere umfassen erfindungsgemäß geeignete Nukleinsäuren wenigstens eine der Teilsequenzen welche für die oben genannten Aminosäuresequenzmotive kodieren.

5 Erfindungsgemäß bevorzugte Nukleinsäuren umfassen Nukleotidsequenzen gemäß SEQ ID NO: 1 und 3, und insbesondere für erfindungsgemäße PARP-Homologe charakteristische Teilsequenzen daraus, wie z.B. Nukleotidsequenzen, umfassend

- 10 a) die Nukleotide +3 bis + 1715 gemäß SEQ ID NO:1;
b) die Nukleotide +242 bis + 1843 gemäß SEQ ID NO:3; oder
c) die Nukleotide +221 bis + 1843 gemäß SEQ ID NO:5;

oder Teilsequenzen von a), b) und c) welche für oben genannte
15 charakteristische Aminosäuresequenzmotive der erfindungsgemäßen PARP-Homologen kodieren.

Ein weiterer Gegenstand der Erfindung umfaßt Expressionskassetten, welche unter genetischer Kontrolle regulativer Nukleotidsequenzen wenigstens eine der oben beschriebenen erfindungsgemäßen
20 Nukleotidsequenzen enthalten. Diese sind zur Herstellung erfindungsgemäßer rekombinanter Vektoren, wie z.B. von viralen Vektoren oder Plasmiden brauchbar, welche wenigstens eine erfindungsgemäße Expressionskassette enthalten.

25 Erfindungsgemäße rekombinante Mikroorganismen sind mit wenigstens einem der oben genannten Vektoren transformiert.

Gegenstand der Erfindung sind außerdem transgene Säuger, transfi-
30 ziert mit einem erfindungsgemäßen Vektor.

Erfindungsgemäß wird außerdem ein in vitro-Screeningverfahren auf Bindungspartner für PARP, insbesondere für ein erfindungsgemäßes PARP-Homologes, bereitgestellt. Eine erste Variante wird so
35 durchgeführt, daß man

- a1) wenigstens ein PARP-Homologes an einem Träger immobilisiert;
b1) das immobilisierte PARP-Homologe mit einem Analyten in Kontakt bringt, in welchem man wenigstens einen Bindungspartner vermutet; und
40 c1) an das immobilisierte PARP-Homologe gebundene Bestandteile des Analyten, gegebenenfalls nach einer Inkubationsphase, bestimmt.

Gemäß einer zweiten Variante wird

- 45 a2) ein Analyt, welcher wenigstens einen möglichen Bindungspartner für das PARP-Homologe enthält, an einem Träger immobilisiert;

- b2) der immobilisierte Analyt mit wenigsten einem PARP-Homologen in Kontakt gebracht, für welches man einen Bindungspartner sucht; und
- c3) der immobilisierte Analyt wird, gegebenenfalls nach einer Inkubationsphase, auf die Bindung des PARP-Homologen untersucht.

Gegenstand der Erfindung ist auch ein Verfahren zur qualitativen oder quantitativen Bestimmung einer PARP-Homologe kodierenden Nukleinsäure, gekennzeichnet durch

- a) Inkubation einer biologischen Probe mit einer definierten Menge einer exogenen erfindungsgemäßen Nukleinsäure (z.B. mit einer Länge von etwa 20 bis 500 Basen oder länger), Hybridisierung, vorzugsweise unter stringenten Bedingungen, Bestimmung der hybridisierenden Nukleinsäuren und gegebenenfalls Vergleich mit einem Standard; oder
- b) Inkubation einer biologischen Probe mit einer definierten Menge von Oligonucleotid-Primerpaaren mit Spezifität für ein PARP-Homologes kodierende Nukleinsäure, Amplifizierung der Nukleinsäure, Bestimmung des Amplifikationsprodukts und gegebenenfalls Vergleich mit einem Standard.

Ein weiterer Gegenstand der Erfindung ist ein Verfahren zur qualitativen oder quantitativen Bestimmung eines erfindungsgemäßen PARP-Homologen, gekennzeichnet durch

- a) Inkubation einer biologischen Probe mit wenigstens einem für ein PARP-Homologes spezifischen Bindungspartner,
- b) Nachweis des Bindungspartner/PARP-Komplexes und gegebenenfalls
- c) Vergleich des Ergebnisses mit einem Standard.

Vorzugsweise ist hierbei der Bindungspartner ein Anti-PARP-Antikörper oder ein bindendes Fragment davon, der gegebenenfalls eine detektierbare Markierung trägt.

Die erfindungsgemäßen Bestimmungsverfahren für PARP, insbesondere für PARP-Homologe und für die kodierenden Nukleinsäuresequenzen davon eignen sich vorteilhafterweise zur Diagnostizierung von Sepsis- oder Ischämie-bedingten Gewebeschädigungen, insbesondere von Schlaganfällen, Myokard-Infarkten oder septischen Schocks.

Weiterhin umfaßt die Erfindung ein Verfahren zur Bestimmung der Wirksamkeit von PARP-Effektoren, das dadurch gekennzeichnet ist, daß man

- a) ein erfindungsgemäßes PARP-Homologes mit einem Analyten inkubiert, welcher einen Effektor einer physiologischen oder pathologischen PARP-Aktivität enthält; den Effektor gegebenenfalls wieder abtrennt; und
- 5 b) die Aktivität des PARP-Homologen, gegebenenfalls nach Zugabe von Substraten oder Cosubstraten, bestimmt.

Ein weiterer Gegenstand der Erfindung betrifft gentherapeutisches Mittel, die in einem gentherapeutisch akzeptablen Träger ein Nukleinsäurekonstrukt enthalten, das

- 10 a) eine Antisense-Nukleinsäure gegen eine kodierende Nukleinsäure gemäß der Erfindung; oder
- b) ein Ribozym gegen eine nicht-kodierende erfindungsgemäße Nukleinsäure umfasst; oder
- 15 c) für einen spezifischen PARP-Inhibitor. kodiert.

Weiterhin betrifft die Erfindung pharmazeutische Mittel, enthaltend in einem pharmazeutisch akzeptablen Träger wenigstens ein

20 erfindungsgemäßes PARP-Protein, wenigstens einen erfindungsgemäßen PARP-Bindungspartner oder wenigstens eine erfindungsgemäße kodierende Nukleotidsequenz.

Schließlich betrifft die Erfindung die Verwendung niedermolekularer (Molekulargewicht von weniger als etwa 1000 Dalton) Bindungspartner eines PARP-Homologen zur Diagnose oder Therapie von

25 Krankheitszuständen, an deren Entstehung und/oder Verlauf wenigstens ein PARP-Protein, insbesondere ein erfindungsgemäßes PARP-Homologes, oder ein davon abgeleitetes Polypeptid beteiligt sind.

30

Die vorliegende Erfindung wird nun unter Bezugnahme auf die beiliegenden Figuren näher beschrieben. Dabei zeigt:

Figur 1 ein Sequenz-Alignment von menschlichem PARP (humanPARP1) und zwei erfindungsgemäß bevorzugte PARPs (humanPARP2 und humanPARP3). Sequenzübereinstimmungen zwischen humanPARP1 und humanPARP2 bzw. humanPARP3 sind umrahmt dargestellt. Die Majoritätssequenz ist über dem Alignment angegeben. Die Zink-Finger Motive von humanPARP1 befinden sich in den Sequenzabschnitten

35

40

chend den Amniosäureresten 21 bis 56 und 125 bis 162;

Figur 2 einen Northern-Blot mit unterschiedlichen menschlichen Geweben zur Veranschaulichung der Gewebeverteilung erfindungsgemäßer PARP-Moleküle. Bahn 1: Herz; Bahn 2: Gehirn; Bahn 3: Plazenta; Bahn 4: Lunge; Bahn 5: Leber; Bahn 6: Skelettmuskel; Bahn 7: Niere; Bahn 8: Pankreas; (A) Blot für humanPARP2; (B) Blot für

45

humanPARP2- und humanPARP3-Sequenzen der Größenstandards (kb) angegeben.

PARP-Homologe und funktionale Äquivalente

5

Soweit keine anderen Angaben gemacht werden, werden im Rahmen der vorliegenden Beschreibung Aminosäuresequenzen beginnend mit dem N-Terminus angegeben. Wird der Ein-Buchstaben-Code für Aminosäuren verwendet, so steht G für Glycin, A für Alanin, V für Valin, L für Leucin, I für Isoleucin, S für Serin, T für Threonin, D für Asparaginsäure, N für Asparagin, E für Glutaminsäure, Q für Glutamin, W für Tryptophan, H für Histidin, R für Arginin, P für Prolin, K für Lysin, Y für Tyrosin, F für Phenylalanin, C für Cystein und M für Methionin.

15

Die vorliegenden Erfindung ist nicht auf die oben konkret beschriebenen PARP-Homologen beschränkt. Vielmehr werden auch solche Homologen erfaßt, welche funktionale Äquivalente davon darstellen. Funktionale Äquivalente umfassen sowohl natürliche, wie

20 z.B. Spezies-spezifische oder Organ-spezifische, als auch künstlich erzeugte Varianten der hierin konkret beschriebenen Proteine. Erfindungsgemäße funktionale Äquivalente unterscheiden sich durch Addition, Substitution, Inversion, Insertion und/oder

Deletion von einem oder mehreren Aminosäureresten von humanPARP2

25 (SEQ ID NO:2) und humanPARP3 (SEQ ID NO: 4 und 6), wobei wenigstens noch die NAD-Bindungsfunktion des Proteins, vermittelt durch eine funktionale katalytische C-terminale Domäne, erhalten bleibt. Funktionale Äquivalente umfassen gegebenenfalls auch solche Varianten, in denen die Leucin-Zipper Region im wesentlichen

30 erhalten bleibt.

Dabei können beispielsweise, ausgehend von der Sequenz für humanPARP2 oder humanPARP3 bestimmte Aminosäuren durch solche mit ähnlichen physikochemischen Eigenschaften (Raumerfüllung, Basizität,

35 Hydrophobizität etc.) ersetzt werden. Beispielsweise können Argininreste gegen Lysinreste, Valinreste gegen Isoleucinreste oder Asparaginsäurereste gegen Glutaminsäurereste ausgetauscht werden. Es können aber auch ein oder mehrere Aminosäuren in ihrer Reihenfolge vertauscht, hinzugefügt oder entfernt werden, oder es können

40 mehrere dieser Maßnahmen miteinander kombiniert werden. Die derart gegenüber der humanPARP2- oder humanPARP3-Sequenz veränderten Proteine besitzen wenigstens 60 %, bevorzugt wenigstens 75 %, ganz besonders bevorzugt wenigstens 85 % Homologie zur Ausgangssequenz, berechnet nach dem Algorithmus von Pearson und Lip-

45 man, Proc. Natl. Acad. Sci (USA) 85(8), 1988, 2444-2448.

Folgende Homologien wurden auf Aminosäureebene bzw. DNA-Ebene zwischen humanPARP1, 2 und 3 bestimmt (FastA-Programm, Pearson und Lipman, a.a.O.):

Aminosäure-Homologien:

5

10

15

	Prozent Identität	Prozent Identität in PARP-Signature
PARP1/PARP2	41,97% (517)	86% (50)
PARP1/PARP3	33,81% (565)	53,1% (49)
PARP2/PARP3	35,20% (537)	53,1% (49)

Zahlen in Klammern geben die Anzahl der überlappenden Aminosäuren an.

DNA-Homologien:

20

25

30

	Prozent Identität im ORF	Prozent Identität in PARP-Signature
PARP1/PARP2	60,81% (467)	77,85% (149)
PARP1/PARP3	58,81% (420)	59,02% (61)
PARP2/PARP3	60,22% (269)	86,36% (22)

Zahlen in Klammern geben die Anzahl der überlappenden Nukleotide an.

Die erfindungsgemäßen Polypeptide lassen sich aufgrund der hohen Ähnlichkeit im Bereich der katalytischen Domäne als homologe

35 Poly(ADP-ribose)polymerasen klassifizieren.

Erfindungswesentlich ist außerdem, daß die neuen PARP-Homologen keine herkömmlichen Zink-Finger Motive aufweisen. Das bedeutet, daß diese Enzyme nicht notwendigerweise in die DNA-Reparatur involviert ist, wohl aber noch ihren pathologischen Mechanismus (NAD⁺-Verbrauch und somit Energieverbrauch durch ATP-Konsum) ausüben können. Dies ist von besonderer Bedeutung für die Wirkstoffentwicklung. Potentielle neue Inhibitoren gegen die erfindungsgemäßen Polymerasen können also die pathologischen Funktionen hemmen ohne negative Effekte auf die gewünschten physiologischen Eigenschaften zu haben. Mit Inhibitoren gegen die bislang bekannte PARPs war dies nicht möglich, da auch immer die DNA-Repa-

raturfunktion mitinhibiert wurde. Die potentiell mutagene Wirkung bekannter PARP-Inhibitoren ist somit leicht verständlich.

Das erfindungsgemäß bevorzugte PARP-Homologe gemäß SEQ ID NO:2
5 (human PARP2) läßt sich vorteilhafterweise aus dem menschlichen Hirn isolieren. In anderen Geweben oder Organen ist humanPARP2 deutlich schwächer exprimiert.

Das erfindungsgemäß bevorzugte PARP-Homologe gemäß SEQ ID NO: 4
10 und 6 (humanPARP3) läßt sich vorteilhafterweise aus dem menschlichen Hirn, Herz oder Niere isolieren. In anderen Geweben oder Organen, wie Muskel oder Leber, ist humanPARP3 deutlich schwächer exprimiert.

15 Der mit der Proteinisolierung vertraute Fachmann wird zur Gewinnung erfindungsgemäßer natürlicher PARPs aus Geweben oder erfindungsgemäßer rekombinant hergestellter PARPs aus Zellkulturen die dazu jeweils am geeignetste Kombination von präparativen Verfahrensmaßnahmen ergreifen. Geeignete präparative Standardmethoden
20 sind zum Beispiel beschrieben in Cooper, T.G., Biochemische Arbeitsmethoden, Verlag Walter de Gruyter, Berlin, New York oder in Scopes, R. Protein Purification, Springer Verlag, New York, Heidelberg, Berlin.

25 Gegenstand der Erfindung sind außerdem PARP2- und PARP3-Homologe, welche zwar aus anderen eukaryotischen Spezies, d.h. Evertebraten oder Vertebraten, insbesondere anderen Säugern, wie z.B. Maus, Ratte, Katze, Hund, Schwein, Schaf, Rind, Pferd, oder Affe oder aus anderen Organen, wie z.B. Myokard, isolierbar sind, aber die
30 wesentlichen, von den erfindungsgemäßen PARPs vorgegebenen strukturellen und funktionellen Eigenschaften besitzen.

Insbesondere das aus menschlichem Hirn isolierbare humanPARP2 und dessen funktionale Äquivalente sind bevorzugte Agenzien für die
35 Entwicklung von Inhibitoren gegen Schlaganfall. Es kann nämlich angenommen werden, daß die Wirkstoffentwicklung, basierend auf PARP2 als Indikator, die Entwicklung von Inhibitoren ermöglicht, welche für die Anwendung an menschlichem Gehirn optimiert sind. Es ist jedoch nicht auszuschließen, daß auf Basis von PARP2 ent-
40 wickelte Inhibitoren auch zur Therapie von PARP-vermittelter pathologischer Zustände anderer Organe ebenfalls einsetzbar sind.

Eine weitere wesentliche biologische Eigenschaft von erfindungsgemäßen PARPs und deren funktionalen Äquivalenten ist in deren
45 Befähigung zur Bindung eines Interaktionspartners zu sehen. Im Unterschied zur bislang bekannten PARPs aus höheren Eukaryoten, wie insbesondere Säugern, verfügen humanPARP2 und 3 über so-

Pflanzenpromotoren CaMV/35S, SSU, OCS, lib4, usp, STLS1, B33, nos oder im Ubiquitin- oder Phaseolin-Promotor enthalten.

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen verwendet werden. Darüberhinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

Die genannten regulatorischen Sequenzen sollen die gezielte Expression der Nukleinsäuresequenzen und der Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, daß das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder daß es sofort exprimiert und/oder überexprimiert wird.

Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Expression positiv beeinflussen und dadurch erhöhen oder erniedrigen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

Unter "Enhancer" sind beispielsweise DNA-Sequenzen zu verstehen, die über eine verbesserte Wechselwirkung zwischen RNA-Polymerase und DNA eine erhöhte Expression bewirken.

Das rekombinante Nukleinsäurekonstrukt bzw. Genkonstrukt wird zur Expression in einem geeigneten Wirtsorganismus vorteilhafterweise in einen wirtsspezifischen Vektor insertiert, der eine optimale Expression der Gene im Wirt ermöglicht. Vektoren sind dem Fachmann wohl bekannt und können beispielsweise aus "Cloning Vectors" (Pouwels P. H. et al., Hrsg, Elsevier, Amsterdam-New York-Oxford, 1985) entnommen werden. Unter Vektoren sind außer Plasmiden auch alle anderen dem Fachmann bekannten Vektoren wie beispielsweise Phagen, Viren, wie SV40, CMV, Baculovirus und Adenovirus, Transposons, IS-Elemente, Phasmide, Cosmide, und lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder chromosomal repliziert werden.

Expression der Konstrukte:

Vorteilhafterweise werden die oben beschriebenen erfindungsgemäßen rekombinanten Konstrukte in ein geeignetes Wirtssystem eingebracht und exprimiert. Dabei werden vorzugsweise dem Fachmann bekannte geläufige Klonierungs- und Transfektionsmethoden verwendet, um die genannten Nukleinsäuren im jeweiligen Expressionssystem

H1

200599

17

stem zur Expression zu bringen. Geeignete Systeme werden beispielsweise in Current Protocols in Molecular Biology, F. Ausubel et al., Hrsg., Wiley Interscience, New York 1997, beschrieben.

5

Als Wirtsorganismen sind prinzipiell alle Organismen geeignet, die eine Expression der erfindungsgemäßen Nukleinsäuren, ihrer Allelvarianten, ihrer funktionellen Äquivalente oder Derivate oder des rekombinanten Nukleinsäurekonstrukts ermöglichen. Unter

- 10 Wirtsorganismen sind beispielsweise Bakterien, Pilze, Hefen, pflanzliche oder tierische Zellen zu verstehen. Bevorzugte Organismen sind Bakterien, wie solche der Gattungen Escherichia, wie z.B. Escherichia coli, Streptomyces, Bacillus oder Pseudomonas, eukaryotische Mikroorganismen, wie Saccharomyces cerevisiae, As-
- 15 pergillus, höhere eukaryotische Zellen aus Tieren oder Pflanzen, beispielsweise Sf9 oder CHO-Zellen.

Gewünschtenfalls kann das Genprodukt auch in transgenen Organismen wie transgenen Tieren, wie insbesondere Mäusen, Schafen oder

- 20 transgenen Pflanzen zur Expression gebracht werden. Bei den transgenen Organismen kann es sich auch um sogenannte Knock-Out Tiere oder Pflanzen handeln, in denen das korrespondierende endogene Gen ausgeschaltet wurde, wie z.B. durch Mutation oder partielle oder vollständige Deletion.

25

Die Kombination aus dem Wirtsorganismen und den zu den Organismen passenden Vektoren, wie Plasmide, Viren oder Phagen, wie beispielsweise Plasmide mit dem RNA-Polymerase/Promoter System, die Phagen λ , μ oder andere temperente Phagen oder Transposons und/

- 30 oder weiteren vorteilhaften regulatorischen Sequenzen bilden ein Expressionssystem. Bevorzugt sind unter dem Begriff Expressionssysteme beispielsweise die Kombination aus Säugetierzellen, wie CHO-Zellen, und Vektoren, wie pcDNA3neo-Vektor, die für Säugetierzellen geeignet sind, zu verstehen.

35

Wie oben beschrieben kann das Genprodukt vorteilhaft auch in transgenen Tieren, z.B. Mäusen, Schafen oder transgenen Pflanzen zur Expression gebracht werden. Ebenso ist es möglich, zellfreie Translationssysteme mit der von der Nukleinsäure abgeleiteten RNA

- 40 zu programmieren.

Darüberhinaus kann das Genprodukt auch in Form therapeutisch oder diagnostisch geeigneter Fragmente exprimiert werden. Zur Isolierung des rekombinanten Proteins können vorteilhaft Vektorsysteme

- 45 oder Oligonukleotide verwendet werden, die die cDNA um bestimmte Nukleotidsequenzen verlängern und damit für veränderte Polypeptide kodieren, die einer einfacheren Reinigung dienen. Derartige

M 29.05.99

18

geeignete Modifikationen sind beispielsweise als Anker fungierende sogenannte "Tags", wie z.B. die als Hexa-Histidin-Anker bekannte Modifikation, oder Epitope, die als Antigene von Antikörpern erkannt werden können (beschrieben zum Beispiel in Harlow, E. and Lane, D., 1988, Antibodies: A Laboratory Manual. Cold Spring Harbor (N.Y.) Press). Diese Anker können zur Anheftung der Proteine an einen festen Träger, wie z.B. einer Polymermatrix, dienen, die beispielsweise in einer Chromatographiesäule eingefüllt sein kann, oder an einer Mikrotiterplatte oder an einem sonstigen Träger verwendet werden kann.

Gleichzeitig können diese Anker auch zur Erkennung der Proteine verwendet werden. Zur Erkennung der Proteine können außerdem übliche Marker, wie Fluoreszenzfarbstoffe, Enzymmarker, die nach Reaktion mit einem Substrat ein detektierbares Reaktionsprodukt bilden, oder radioaktive Marker, allein oder in Kombination mit den Ankern zur Derivatisierung der Proteine verwendet werden.

Herstellung von Antikörpern:

20

Die Herstellung von Anti-PARP2-Antikörpern erfolgt in einer dem Fachmann geläufigen Weise. Mit Antikörpern sind sowohl polyklonale, monoklonale, humane oder humanisierte Antikörper oder Fragmente davon, single chain Antikörper oder auch synthetische Antikörper gemeint, ebenso wie Antikörperfragmente, wie Fv, Fab und (Fab)'₂. Geeignete Herstellungsverfahren sind z.B. beschrieben in Campbell, A.M., Monoclonal Antibody Technology, (1987) Elsevier Verlag, Amsterdam, New York, Oxford und in Breitling, F. und Dübel, S., Rekombinante Antikörper (1997), Spektrum Akademischer Verlag, Heidelberg.

Weitere Anwendung der kodierenden Sequenz:

Die vorliegende cDNA bietet außerdem die Voraussetzung, die genomische Sequenz des neuen PARP-Gens zu klonieren. Darunter fällt auch die dazugehörige regulatorische oder Promotorsequenz, die beispielsweise durch Sequenzierung des 5'-stromaufwärts gelegenen Bereiches der erfindungsgemäßen cDNA zugänglich wird. Die Sequenzinformation der cDNA ist auch die Grundlage für die Herstellung von Antisense-Molekülen oder Ribozymen mit Hilfe bekannter Methoden (vgl. Jones, J.T. und Sallenger, B.A. (1997) Nat. Biotechnol. 15, 902; Nellen, W. und Lichtenstein, C. (1993) TIBS, 18, 419). Die genomische DNA kann ebenfalls zur Herstellung der oben beschriebenen Genkonstrukte verwendet werden.

45

Eine weitere Möglichkeit des Einsatzes der Nukleotidsequenz oder Teilen davon ist die Erzeugung transgener Tiere. Transgene Überexpression oder genetischer Knockout der Sequenzinformation in geeigneten Tiermodellen kann wertvolle weitere Informationen über die (Patho-)Physiologie der neuen Enzyme liefern.

Therapeutische Anwendungen:

In Situationen, in denen ein Mangel an einem erfindungsgemäßen Protein herrscht, können mehrere Methoden zur Substituierung eingesetzt werden. Zum einen kann das Protein, natürlich oder rekombinant direkt, oder gentherapeutisch in Form seiner kodierenden Nukleinsäure (DNA oder RNA) appliziert werden. Dazu können beliebige Vektoren beispielsweise sowohl virale, als auch nichtvirale Vehikel zum Einsatz kommen. Geeignete Methoden werden beispielsweise beschrieben von Strauss und Barranger in Concepts in Gene Therapy (1997), Walter de Gruyter, Hrsg. Eine weitere Alternative bietet die Stimulation des endogenen, körpereigenen Genes durch geeignete Mittel.

20

Auch der turn-over oder die Inaktivierung erfindungsgemäßer PARPs, z.B. durch Proteasen, können blockiert werden. Schließlich können Inhibitoren oder Agonisten erfindungsgemäßer PARPs zum Einsatz gelangen.

25

In Situationen, in denen überschüssiges PARP vorliegt, können unterschiedliche Typen von Inhibitoren eingesetzt werden. Diese Inhibition kann sowohl durch Antisense-Moleküle, Ribozyme, Oligonukleotide oder Antikörper, als auch durch niedermolekulare Verbindungen erreicht werden.

30

Nicht-therapeutische Anwendungen:

Weiterhin können die erfindungsgemäßen Nukleinsäuren, wie z.B. cDNA, die genomische DNA, der Promotor, als auch das Polypeptid, sowie Teilfragmente davon in rekombinanter oder nichtrekombinanter Form zur Ausarbeitung von verschiedenen Testsystemen verwendet werden.

Beispielsweise kann ein Testsystem etabliert werden, das geeignet ist, die Aktivität des Promotors oder des Proteins in Anwesenheit einer Testsubstanz zu messen. Bevorzugt handelt es sich dabei um einfache, z.B. colorimetrische, luminometrische, fluorimetrische, immunologische oder radioaktive, Meßmethoden, die die schnelle Meßbarkeit, vorzugsweise einer Vielzahl von Testsubstanzen erlauben. Derartige Tests eignen sich vorteilhaft für ein sogenanntes High-Throughput-Screening. Diese Testsysteme erlauben eine Bewer-

40

45

M

29.05.99

tung von Testsubstanzen in Bezug auf deren Bindung an oder deren Agonisierung, Antagonisierung oder Inhibition von erfindungsgemäßen Proteinen.

- 5 Die Bestimmung von Menge, Aktivität und Verteilung der erfindungsgemäßen Proteine oder ihrer zugrundeliegenden mRNA im menschlichen Körper kann zur Diagnose, zur Bestimmung der Prädisposition und zum Monitoring bei bestimmten Erkrankungen dienen. Desgleichen kann die Sequenz der cDNA sowie der genomischen Sequenz zu Aussagen über genetische Ursachen und Prädispositionen bestimmter Erkrankungen herangezogen werden. Dazu können sowohl DNA/RNA-Sonden als auch Antikörper verschiedenster Art benutzt werden. Weiterhin können die erfindungsgemäßen Nukleotidsequenzen oder Teile davon in Form geeigneter Sonden zur Aufdeckung von Punktmutationen, Deletionen oder Insertionen dienen.

- Weiterhin können die erfindungsgemäßen Proteine benutzt werden, um ihre natürlichen Liganden oder Interaktionspartner zu bestimmen und zu isolieren. Außerdem können die erfindungsgemäßen Proteine benutzt werden, um künstliche oder synthetische Liganden zu bestimmen und zu isolieren. Dazu kann man das rekombinant hergestellte oder gereinigte natürliche Protein derart derivatisieren, daß es Modifikationen trägt, die eine Verknüpfung mit Trägermaterialien erlauben. Derart gebundenen Proteine können mit unterschiedlichen Analyten, wie z.B. Proteinextrakten oder Peptidbibliotheken oder anderen Quellen für Liganden, inkubiert werden. Spezifisch gebundenen Peptide, Proteine oder niedermolekulare, nichtproteinogene Substanzen können so isoliert und charakterisiert werden. Unter nichtproteinogenen Substanzen sind beispielsweise niedermolekulare, chemische Substanzen (= kleiner 1000 Dalton) zu verstehen, die beispielsweise aus der klassischen Wirkstoffsynthese oder aus sogenannten Substanzbibliotheken stammen können, die mit Hilfe der Kombinatorik synthetisiert worden sind.

- 35 Die verwendeten Proteinextrakte sind beispielsweise abgeleitet aus Homogenaten von Pflanzen oder Pflanzenteilen, Mikroorganismen, menschlichen oder tierischen Geweben oder Organen.

- Liganden oder Interaktionspartner können ferner durch Verfahren, wie das Hefe "Two-Hybrid-System" identifiziert werden (Fields, S. und Song, O. (1989) Nature, 340, 245). Die dabei einsetzbaren Expressionsbanken sind beispielsweise ableitbar aus menschlichen Geweben, wie z.B. Gehirn, Herz, Niere usw.

- 45 Die erfindungsgemäßen Nukleinsäuresequenzen und die von ihnen kodierten Proteine können zur Entwicklung von Reagenzien, Agonisten und Antagonisten oder Inhibitoren zur Diagnose und Therapie von

M 2.05.99

21

chronischen und akuten Erkrankungen, die mit der Expression einer der erfindungsgemäßen Proteinsequenzen, wie z.B. mit deren erhöhter oder erniedrigter Expression assoziiert sind, eingesetzt werden. Die entwickelten Reagenzien, Agonisten, Antagonisten oder

5 Inhibitoren können anschließend zur Herstellung von pharmazeutischen Zubereitungen zur Behandlung oder Diagnose von Krankheiten verwendet. Dabei kann es sich beispielsweise um Erkrankungen des Gehirns, des Herz-Kreislaufsystems oder des Auges, oder septischem Schock handeln.

10

Die Erfindung wird nun unter Bezugnahme auf die folgenden Ausführungsbeispiele näher erläutert.

Beispiel 1: Isolierung der PARP2- und PARP3-cDNA

15

Bei der Sequenzanalyse von cDNA-Klonen einer cDNA-Bibliothek aus menschlichem Gehirn (Human Brain 5' Stretch Plus cDNA Library, # HL3002a, Fa. Clontech) wurden die vorliegenden cDNA-Sequenzen erstmalig gefunden. Die Sequenzen dieser Klone sind in SEQ ID

20 NO:1 und 3 beschrieben.

Beispiel 2: Expression von humanPARP2 bzw. humanPARP3 in menschlichen Geweben

25 Die Expression von humanPARP2 bzw. humanPARP3 wurde in acht verschiedenen menschlichen Geweben mittels Northern-Blot-Analyse untersucht. Ein "Human Multiple Tissue Northern Blot" der Firma Clontech (#7760-1) wurde dazu mit einer RNA-Sonde hybridisiert. Die Sonde wurde durch *in vitro* Transkription der entsprechenden

30 cDNA von humanem PARP2 bzw. humanem PARP3 in Anwesenheit Digoxigenin-markierter Nukleotide hergestellt.

Nach stringenter Waschung wurde das Transkript von humanem PARP2 hauptsächlich im Gehirn nachgewiesen, leicht ist es ebenfalls im

35 Herzen exprimiert. Die Expression in weiteren Geweben (Plazenta, Lunge, Leber, Skelettmuskel, Niere, Bauchspeicheldrüse) ist sehr schwach. Die Größe des Transkriptes von ca. 1.9 kb entspricht der Länge der bestimmten cDNA (1.85kb) (vgl. Figur 2(A)).

40 Nach stringenter Waschung wurde das Transkript von humanem PARP3 hauptsächlich im Herzen, Gehirn und Niere nachgewiesen, deutlich aber schwächer ist es ebenfalls in Skelettmuskel und Leber exprimiert. Die Expression in weiteren Geweben (Plazenta, Lunge, Bauchspeicheldrüse) ist deutlich schwächer (vgl. Figur 2(B)). Zu

45 humanPARP3 gibt es mindestens 2 Transkripte. Deren Größe (ca. 2,2 kb bzw. 2,5 kb) entspricht der Länge der bestimmten cDNA (2.3kb).

M

29.05.99

22

Für die stringente Waschung wurde ein 0,1X SSC Puffer (hergestellt aus 20X SSC: 3M NaCl, 0,3M Natriumcitrat, pH 7,0), supplementiert mit 0,1% SDS bei 68°C verwendet.

5

10

15

20

25

30

35

40

45

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) ANMELDER:

- (A) NAME: BASF Aktiengesellschaft
- (B) STRASSE: -
- (C) ORT: Ludwigshafen
- (E) LAND: Deutschland
- (F) POSTLEITZAHL: 67065

(ii) BEZEICHNUNG DER ERFINDUNG: Neue Poly-ADP-Ribose-Polymerase Gene

(iii) ANZAHL DER SEQUENZEN: 6

(iv) COMPUTER-LESBARE FASSUNG:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) ANGABEN ZU SEQ ID NO: 1:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1843 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

(F) GEWEBETYP: Brain

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE: 3..1715
- (D) SONSTIGE ANGABEN: /product= "Poly ADP Ribose Polymerase"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

CC ATG GCG GCG CGG CGG CGA CGG AGC ACC GGC GGC GGC AGG GCG AGA
Met Ala Ala Arg Arg Arg Arg Ser Thr Gly Gly Gly Arg Ala Arg

47

24

H

Z

0

5

0

1

5

10

15

GCA	TTA	AAT	GAA	AGC	AAA	AGA	GTT	AAT	AAT	GGC	AAC	ACG	GCT	CCA	GAA	95
Ala	Leu	Asn	Glu	Ser	Lys	Arg	Val	Asn	Asn	Gly	Asn	Thr	Ala	Pro	Glu	
					20				25					30		
GAC	TCT	TCC	CCT	GCC	AAG	AAA	ACT	CGT	AGA	TGC	CAG	AGA	CAG	GAG	TCG	143
Asp	Ser	Ser	Pro	Ala	Lys	Lys	Thr	Arg	Arg	Cys	Gln	Arg	Gln	Glu	Ser	
			35					40					45			
AAA	AAG	ATG	CCT	GTG	GCT	GGA	GGA	AAA	GCT	AAT	AAG	GAC	AGG	ACA	GAA	191
Lys	Lys	Met	Pro	Val	Ala	Gly	Gly	Lys	Ala	Asn	Lys	Asp	Arg	Thr	Glu	
		50					55					60				
GAC	AAG	CAA	GAT	GAA	TCT	GTG	AAG	GCC	TTG	CTG	TTA	AAG	GGC	AAA	GCT	239
Asp	Lys	Gln	Asp	Glu	Ser	Val	Lys	Ala	Leu	Leu	Leu	Lys	Gly	Lys	Ala	
	65					70					75					
CCT	GTG	GAC	CCA	GAG	TGT	ACA	GCC	AAG	GTG	GGG	AAG	GCT	CAT	GTG	TAT	287
Pro	Val	Asp	Pro	Glu	Cys	Thr	Ala	Lys	Val	Gly	Lys	Ala	His	Val	Tyr	
	80					85				90					95	
TGT	GAA	GGA	AAT	GAT	GTC	TAT	GAT	GTC	ATG	CTA	AAT	CAG	ACC	AAT	CTC	335
Cys	Glu	Gly	Asn	Asp	Val	Tyr	Asp	Val	Met	Leu	Asn	Gln	Thr	Asn	Leu	
				100					105					110		
CAG	TTC	AAC	AAC	AAC	AAG	TAC	TAT	CTG	ATT	CAG	CTA	TTA	GAA	GAT	GAT	383
Gln	Phe	Asn	Asn	Asn	Lys	Tyr	Tyr	Leu	Ile	Gln	Leu	Leu	Glu	Asp	Asp	
			115					120					125			
GCC	CAG	AGG	AAC	TTC	AGT	GTT	TGG	ATG	AGA	TGG	GGC	CGA	GTT	GGG	AAA	431
Ala	Gln	Arg	Asn	Phe	Ser	Val	Trp	Met	Arg	Trp	Gly	Arg	Val	Gly	Lys	
		130					135					140				
ATG	GGA	CAG	CAC	AGC	CTG	GTG	GCT	TGT	TCA	GGC	AAT	CTC	AAC	AAG	GCC	479
Met	Gly	Gln	His	Ser	Leu	Val	Ala	Cys	Ser	Gly	Asn	Leu	Asn	Lys	Ala	
	145					150					155					
AAG	GAA	ATC	TTT	CAG	AAG	AAA	TTC	CTT	GAC	AAA	ACG	AAA	AAC	AAT	TGG	527
Lys	Glu	Ile	Phe	Gln	Lys	Lys	Phe	Leu	Asp	Lys	Thr	Lys	Asn	Asn	Trp	
	160					165				170					175	
GAA	GAT	CGA	GAA	AAG	TTT	GAG	AAG	GTG	CCT	GGA	AAA	TAT	GAT	ATG	CTA	575
Glu	Asp	Arg	Glu	Lys	Phe	Glu	Lys	Val	Pro	Gly	Lys	Tyr	Asp	Met	Leu	
				180					185					190		
CAG	ATG	GAC	TAT	GCC	ACC	AAT	ACT	CAG	GAT	GAA	GAG	GAA	ACA	AAG	AAA	623

25

Gln Met Asp Tyr Ala Thr Asn Thr Gln Asp Glu Glu Glu Thr Lys Lys
 195 200 205

GAG GAA TCT CTT AAA TCT CCC TTG AAG CCA GAG TCA CAG CTA GAT CTT 671
 Glu Glu Ser Leu Lys Ser Pro Leu Lys Pro Glu Ser Gln Leu Asp Leu
 210 215 220

CGG GTA CAG GAG TTA ATA AAG TTG ATC TGT AAT GTT CAG GCC ATG GAA 719
 Arg Val Gln Glu Leu Ile Lys Leu Ile Cys Asn Val Gln Ala Met Glu
 225 230 235

GAA ATG ATG ATG GAA ATG AAG TAT AAT ACC AAG AAA GCC CCA CTT GGG 767
 Glu Met Met Met Glu Met Lys Tyr Asn Thr Lys Lys Ala Pro Leu Gly
 240 245 250 255

AAG CTG ACA GTG GCA CAA ATC AAG GCA GGT TAC CAG TCT CTT AAG AAG 815
 Lys Leu Thr Val Ala Gln Ile Lys Ala Gly Tyr Gln Ser Leu Lys Lys
 260 265 270

ATT GAG GAT TGT ATT CGG GCT GGC CAG CAT GGA CGA GCT CTC ATG GAA 863
 Ile Glu Asp Cys Ile Arg Ala Gly Gln His Gly Arg Ala Leu Met Glu
 275 280 285

GCA TGC AAT GAA TTC TAC ACC AGG ATT CCG CAT GAC TTT GGA CTC CGT 911
 Ala Cys Asn Glu Phe Tyr Thr Arg Ile Pro His Asp Phe Gly Leu Arg
 290 295 300

ACT CCT CCA CTA ATC CGG ACA CAG AAG GAA CTG TCA GAA AAA ATA CAA 959
 Thr Pro Pro Leu Ile Arg Thr Gln Lys Glu Leu Ser Glu Lys Ile Gln
 305 310 315

TTA CTA GAG GCT TTG GGA GAC ATT GAA ATT GCT ATT AAG CTG GTG AAA 1007
 Leu Leu Glu Ala Leu Gly Asp Ile Glu Ile Ala Ile Lys Leu Val Lys
 320 325 330 335

ACA GAG CTA CAA AGC CCA GAA CAC CCA TTG GAC CAA CAC TAT AGA AAC 1055
 Thr Glu Leu Gln Ser Pro Glu His Pro Leu Asp Gln His Tyr Arg Asn
 340 345 350

CTA CAT TGT GCC TTG CGC CCC CTT GAC CAT GAA AGT TAC GAG TTC AAA 1103
 Leu His Cys Ala Leu Arg Pro Leu Asp His Glu Ser Tyr Glu Phe Lys
 355 360 365

GTG ATT TCC CAG TAC CTA CAA TCT ACC CAT GCT CCC ACA CAC AGC GAC 1151
 Val Ile Ser Gln Tyr Leu Gln Ser Thr His Ala Pro Thr His Ser Asp
 370 375 380

26

TAT	ACC	ATG	ACC	TTG	CTG	GAT	TTG	TTT	GAA	GTG	GAG	AAG	GAT	GGT	GAG	1199
Tyr	Thr	Met	Thr	Leu	Leu	Asp	Leu	Phe	Glu	Val	Glu	Lys	Asp	Gly	Glu	
	385					390					395					
AAA	GAA	GCC	TTC	AGA	GAG	GAC	CTT	CAT	AAC	AGG	ATG	CTT	CTA	TGG	CAT	1247
Lys	Glu	Ala	Phe	Arg	Glu	Asp	Leu	His	Asn	Arg	Met	Leu	Leu	Trp	His	
400					405					410					415	
GGT	TCC	AGG	ATG	AGT	AAC	TGG	GTG	GGA	ATC	TTG	AGC	CAT	GGG	CTT	CGA	1295
Gly	Ser	Arg	Met	Ser	Asn	Trp	Val	Gly	Ile	Leu	Ser	His	Gly	Leu	Arg	
				420					425					430		
ATT	GCC	CCA	CCT	GAA	GCT	CCC	ATC	ACA	GGT	TAC	ATG	TTT	GGG	AAA	GGA	1343
Ile	Ala	Pro	Pro	Glu	Ala	Pro	Ile	Thr	Gly	Tyr	Met	Phe	Gly	Lys	Gly	
			435					440					445			
ATC	TAC	TTT	GCT	GAC	ATG	TCT	TCC	AAG	AGT	GCC	AAT	TAC	TGC	TTT	GCC	1391
Ile	Tyr	Phe	Ala	Asp	Met	Ser	Ser	Lys	Ser	Ala	Asn	Tyr	Cys	Phe	Ala	
		450					455					460				
TCT	CGC	CTA	AAG	AAT	ACA	GGA	CTG	CTG	CTC	TTA	TCA	GAG	GTA	GCT	CTA	1439
Ser	Arg	Leu	Lys	Asn	Thr	Gly	Leu	Leu	Leu	Leu	Ser	Glu	Val	Ala	Leu	
	465					470					475					
GGT	CAG	TGT	AAT	GAA	CTA	CTA	GAG	GCC	AAT	CCT	AAG	GCC	GAA	GGA	TTG	1487
Gly	Gln	Cys	Asn	Glu	Leu	Leu	Glu	Ala	Asn	Pro	Lys	Ala	Glu	Gly	Leu	
480					485					490					495	
CTT	CAA	GGT	AAA	CAT	AGC	ACC	AAG	GGG	CTG	GGC	AAG	ATG	GCT	CCC	AGT	1535
Leu	Gln	Gly	Lys	His	Ser	Thr	Lys	Gly	Leu	Gly	Lys	Met	Ala	Pro	Ser	
				500					505					510		
TCT	GCC	CAC	TTC	GTC	ACC	CTG	AAT	GGG	AGT	ACA	GTG	CCA	TTA	GGA	CCA	1583
Ser	Ala	His	Phe	Val	Thr	Leu	Asn	Gly	Ser	Thr	Val	Pro	Leu	Gly	Pro	
			515					520					525			
GCA	AGT	GAC	ACA	GGA	ATT	CTG	AAT	CCA	GAT	GGT	TAT	ACC	CTC	AAC	TAC	1631
Ala	Ser	Asp	Thr	Gly	Ile	Leu	Asn	Pro	Asp	Gly	Tyr	Thr	Leu	Asn	Tyr	
		530					535					540				
AAT	GAA	TAT	ATT	GTA	TAT	AAC	CCC	AAC	CAG	GTC	CGT	ATG	CGG	TAC	CTT	1679
Asn	Glu	Tyr	Ile	Val	Tyr	Asn	Pro	Asn	Gln	Val	Arg	Met	Arg	Tyr	Leu	
	545					550					555					
TTA	AAG	GTT	CAG	TTT	AAT	TTC	CTT	CAG	CTG	TGG	TGA	ATGTTGATAT				1725
Leu	Lys	Val	Gln	Phe	Asn	Phe	Leu	Gln	Leu	Trp	*					

TAAATAAACC AGAGATCTGA TCTTCAAGCA AGAAAATAAG CAGTGGTTGTA CTTGTGAATT

1785

TTGTGATATT TTATGTAATA AAAACTGTAC AGGTCTAAAA AAAAAAAAAA AAAAAAAAAA

1843

(2) ANGABEN ZU SEQ ID NO: 2:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 571 Aminosäuren

(B) ART: Aminosäure

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

Met Ala Ala Arg Arg Arg Arg Ser Thr Gly Gly Gly Arg Ala Arg Ala
1 5 10 15

Leu Asn Glu Ser Lys Arg Val Asn Asn Gly Asn Thr Ala Pro Glu Asp
20 25 30

Ser Ser Pro Ala Lys Lys Thr Arg Arg Cys Gln Arg Gln Glu Ser Lys
35 40 45

Lys Met Pro Val Ala Gly Gly Lys Ala Asn Lys Asp Arg Thr Glu Asp
50 55 60

Lys Gln Asp Glu Ser Val Lys Ala Leu Leu Leu Lys Gly Lys Ala Pro
65 70 75 80

Val Asp Pro Glu Cys Thr Ala Lys Val Gly Lys Ala His Val Tyr Cys
85 90 95

Glu Gly Asn Asp Val Tyr Asp Val Met Leu Asn Gln Thr Asn Leu Gln
100 105 110

Phe Asn Asn Asn Lys Tyr Tyr Leu Ile Gln Leu Leu Glu Asp Asp Ala
115 120 125

Gln Arg Asn Phe Ser Val Trp Met Arg Trp Gly Arg Val Gly Lys Met
130 135 140

Gly Gln His Ser Leu Val Ala Cys Ser Gly Asn Leu Asn Lys Ala Lys
145 150 155 160

Glu Ile Phe Gln Lys Lys Phe Leu Asp Lys Thr Lys Asn Asn Trp Glu

28

165

170

175

Asp Arg Glu Lys Phe Glu Lys Val Pro Gly Lys Tyr Asp Met Leu Gln
180 185 190

Met Asp Tyr Ala Thr Asn Thr Gln Asp Glu Glu Glu Thr Lys Lys Glu
195 200 205

Glu Ser Leu Lys Ser Pro Leu Lys Pro Glu Ser Gln Leu Asp Leu Arg
210 215 220

Val Gln Glu Leu Ile Lys Leu Ile Cys Asn Val Gln Ala Met Glu Glu
225 230 235 240

Met Met Met Glu Met Lys Tyr Asn Thr Lys Lys Ala Pro Leu Gly Lys
245 250 255

Leu Thr Val Ala Gln Ile Lys Ala Gly Tyr Gln Ser Leu Lys Lys Ile
260 265 270

Glu Asp Cys Ile Arg Ala Gly Gln His Gly Arg Ala Leu Met Glu Ala
275 280 285

Cys Asn Glu Phe Tyr Thr Arg Ile Pro His Asp Phe Gly Leu Arg Thr
290 295 300

Pro Pro Leu Ile Arg Thr Gln Lys Glu Leu Ser Glu Lys Ile Gln Leu
305 310 315 320

Leu Glu Ala Leu Gly Asp Ile Glu Ile Ala Ile Lys Leu Val Lys Thr
325 330 335

Glu Leu Gln Ser Pro Glu His Pro Leu Asp Gln His Tyr Arg Asn Leu
340 345 350

His Cys Ala Leu Arg Pro Leu Asp His Glu Ser Tyr Glu Phe Lys Val
355 360 365

Ile Ser Gln Tyr Leu Gln Ser Thr His Ala Pro Thr His Ser Asp Tyr
370 375 380

Thr Met Thr Leu Leu Asp Leu Phe Glu Val Glu Lys Asp Gly Glu Lys
385 390 395 400

Glu Ala Phe Arg Glu Asp Leu His Asn Arg Met Leu Leu Trp His Gly
405 410 415

29

Ser Arg Met Ser Asn Trp Val Gly Ile Leu Ser His Gly Leu Arg Ile
420 425 430

Ala Pro Pro Glu Ala Pro Ile Thr Gly Tyr Met Phe Gly Lys Gly Ile
435 440 445

Tyr Phe Ala Asp Met Ser Ser Lys Ser Ala Asn Tyr Cys Phe Ala Ser
450 455 460

Arg Leu Lys Asn Thr Gly Leu Leu Leu Leu Ser Glu Val Ala Leu Gly
465 470 475 480

Gln Cys Asn Glu Leu Leu Glu Ala Asn Pro Lys Ala Glu Gly Leu Leu
485 490 495

Gln Gly Lys His Ser Thr Lys Gly Leu Gly Lys Met Ala Pro Ser Ser
500 505 510

Ala His Phe Val Thr Leu Asn Gly Ser Thr Val Pro Leu Gly Pro Ala
515 520 525

Ser Asp Thr Gly Ile Leu Asn Pro Asp Gly Tyr Thr Leu Asn Tyr Asn
530 535 540

Glu Tyr Ile Val Tyr Asn Pro Asn Gln Val Arg Met Arg Tyr Leu Leu
545 550 555 560

Lys Val Gln Phe Asn Phe Leu Gln Leu Trp *
565 570

(2) ANGABEN ZU SEQ ID NO: 3:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 2265 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

- (F) GEWEBETYP: Uterus

M

29.05.99

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) LAGE:242..1843

(D) SONSTIGE ANGABEN:/product= "Poly ADP Ribose
Polymerase"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

TGGGACTGGT CGCCTGACTC GGCCTGCCCC AGCCTCTGCT TCACCCCACT GGTGGCCAAA	60
TAGCCGATGT CTAATCCCC ACACAAGCTC ATCCCCGGCC TCTGGGATTG TTGGGAATTC	120
TCTCCCTAAT TCACGCCTGA GGCTCATGGA GAGTTGCTAG ACCTGGGACT GCCCTGGGAG	180
GCGCACACAA CCAGGCCGGG TGGCAGCCAG GACCTCTCCC ATGTCCCTGC TTTTCTTGGC	240
C ATG GCT CCA AAG CCG AAG CCC TGG GTA CAG ACT GAG GGC CCT GAG	286
Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro Glu	
575 580 585	
AAG AAG AAG GGC CGG CAG GCA GGA AGG GAG GAG GAC CCC TTC CGC TCC	334
Lys Lys Lys Gly Arg Gln Ala Gly Arg Glu Glu Asp Pro Phe Arg Ser	
590 595 600	
ACC GCT GAG GCC CTC AAG GCC ATA CCC GCA GAG AAG CGC ATA ATC CGC	382
Thr Ala Glu Ala Leu Lys Ala Ile Pro Ala Glu Lys Arg Ile Ile Arg	
605 610 615	
GTG GAT CCA ACA TGT CCA CTC AGC AGC AAC CCC GGG ACC CAG GTG TAT	430
Val Asp Pro Thr Cys Pro Leu Ser Ser Asn Pro Gly Thr Gln Val Tyr	
620 625 630	
GAG GAC TAC AAC TGC ACC CTG AAC CAG ACC AAC ATC GAG AAC AAC AAC	478
Glu Asp Tyr Asn Cys Thr Leu Asn Gln Thr Asn Ile Glu Asn Asn Asn	
635 640 645 650	
AAC AAG TTC TAC ATC ATC CAG CTG CTC CAA GAC AGC AAC CGC TTC TTC	526
Asn Lys Phe Tyr Ile Ile Gln Leu Leu Gln Asp Ser Asn Arg Phe Phe	
655 660 665	
ACC TGC TGG AAC CGC TGG GGC CGT GTG GGA GAG GTC GGC CAG TCA AAG	574
Thr Cys Trp Asn Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys	
670 675 680	
ATC AAC CAC TTC ACA AGG CTA GAA GAT GCA AAG AAG GAC TTT GAG AAG	622
Ile Asn His Phe Thr Arg Leu Glu Asp Ala Lys Lys Asp Phe Glu Lys	

31

H

29.05.99

685

690

695

AAA TTT CGG GAA AAG ACC AAG AAC AAC TGG GCA GAG CGG GAC CAC TTT	670
Lys Phe Arg Glu Lys Thr Lys Asn Asn Trp Ala Glu Arg Asp His Phe	
700 705 710	
GTG TCT CAC CCG GGC AAG TAC ACA CTT ATC GAA GTA CAG GCA GAG GAT	718
Val Ser His Pro Gly Lys Tyr Thr Leu Ile Glu Val Gln Ala Glu Asp	
715 720 725 730	
GAG GCC CAG GAA GCT GTG GTG AAG GTG GAC AGA GGC CCA GTG AGG ACT	766
Glu Ala Gln Glu Ala Val Val Lys Val Asp Arg Gly Pro Val Arg Thr	
735 740 745	
GTG ACT AAG CGG GTG CAG CCC TGC TCC CTG GAC CCA GCC ACG CAG AAG	814
Val Thr Lys Arg Val Gln Pro Cys Ser Leu Asp Pro Ala Thr Gln Lys	
750 755 760	
CTC ATC ACT AAC ATC TTC AGC AAG GAG ATG TTC AAG AAC ACC ATG GCC	862
Leu Ile Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Thr Met Ala	
765 770 775	
CTC ATG GAC CTG GAT GTG AAG AAG ATG CCC CTG GGA AAG CTG AGC AAG	910
Leu Met Asp Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Ser Lys	
780 785 790	
CAA CAG ATT GCA CGG GGT TTC GAG GCC TTG GAG GCG CTG GAG GAG GCC	958
Gln Gln Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala	
795 800 805 810	
CTG AAA GGC CCC ACG GAT GGT GGC CAA AGC CTG GAG GAG CTG TCC TCA	1006
Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser Ser	
815 820 825	
CAC TTT TAC ACC GTC ATC CCG CAC AAC TTC GGC CAC AGC CAG CCC CCG	1054
His Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro Pro	
830 835 840	
CCC ATC AAT TCC CCT GAG CTT CTG CAG GCC AAG AAG GAC ATG CTG CTG	1102
Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu Leu	
845 850 855	
GTG CTG GCG GAC ATC GAG CTG GCC CAG GCC CTG CAG GCA GTC TCT GAG	1150
Val Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser Glu	
860 865 870	
CAG GAG AAG ACG GTG GAG GAG GTG CCA CAC CCC CTG GAC CGA GAC TAC	1198

32

H

2005-99

Gln Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp Tyr
 875 880 885 890

CAG CTT CTC AAG TGC CAG CTG CAG CTG CTA GAC TCT GGA GCA CCT GAG 1246
 Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro Glu
 895 900 905

TAC AAG GTG ATA CAG ACC TAC TTA GAA CAG ACT GGC AGC AAC CAC AGG 1294
 Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His Arg
 910 915 920

TGC CCT ACA CTT CAA CAC ATC TGG AAA GTA AAC CAA GAA GGG GAG GAA 1342
 Cys Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu Glu
 925 930 935

GAC AGA TTC CAG GCC CAC TCC AAA CTG GGT AAT CGG AAG CTG CTG TGG 1390
 Asp Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu Trp
 940 945 950

CAT GGC ACC AAC ATG GCC GTG GTG GCC GCC ATC CTC ACT AGT GGG CTC 1438
 His Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly Leu
 955 960 965 970

CGC ATC ATG CCA CAT TCT GGT GGG CGT GTT GGC AAG GGC ATC TAC TTT 1486
 Arg Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe
 975 980 985

GCC TCA GAG AAC AGC AAG TCA GCT GGA TAT GTT ATT GGC ATG AAG TGT 1534
 Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys Cys
 990 995 1000

GGG GCC CAC CAT GTC GGC TAC ATG TTC CTG GGT GAG GTG GCC CTG GGC 1582
 Gly Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly
 1005 1010 1015

AGA GAG CAC CAT ATC AAC ACG GAC AAC CCC AGC TTG AAG AGC CCA CCT 1630
 Arg Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro Pro
 1020 1025 1030

CCT GGC TTC GAC AGT GTC ATT GCC CGA GGC CAC ACC GAG CCT GAT CCG 1678
 Pro Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp Pro
 1035 1040 1045 1050

ACC CAG GAC ACT GAG TTG GAG CTG GAT GGC CAG CAA GTG GTG GTG CCC 1726
 Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val Pro
 1055 1060 1065

33

H

29.05.99

CAG GGC CAG CCT GTG CCC TGC CCA GAG TTC AGC AGC TCC ACA TTC TCC 1774
Gln Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser
1070 1075 1080

CAG AGC GAG TAC CTC ATC TAC CAG GAG AGC CAG TGT CGC CTG CGC TAC 1822
Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr
1085 1090 1095

CTG CTG GAG GTC CAC CTC TGA GTGCCCCGCC TGTCCCCCGG GGTCTGCAA 1873
Leu Leu Glu Val His Leu *
1100 1105

GGCTGGACTG TGATCTTCAA TCATCCTGCC CATCTCTGGT ACCCCTATAT CACTCCTTTT 1933

TTTCAAGAAT ACAATACGTT GTTGTTAACT ATAGTCACCA TGCTGTACAA GATCCCTGAA 1993

CTTATGCCTC CTAAGTAAAA TTTTGTATTC TTTGACACAT CTGCCCAGTC CCTCTCCTCC 2053

CAGCCCATGG TAACCAGCAT TTGACTCTTT ACTTGTATAA GGGCAGCTTT TATAGGTTCC 2113

ACATGTAAGT GAGATCATGC AGTGTTTGTC TTTCTGTGCC TGGCTTATTT CACTCAGCAT 2173

AATGTGCACC GGGTTCACCC ATGTTTTCAT AAATGACAAG ATTTCTCCTT TAAAAA AAAA 2233

AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 2265

(2) ANGABEN ZU SEQ ID NO: 4:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 534 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro Glu Lys
1 5 10 15

Lys Lys Gly Arg Gln Ala Gly Arg Glu Glu Asp Pro Phe Arg Ser Thr
20 25 30

Ala Glu Ala Leu Lys Ala Ile Pro Ala Glu Lys Arg Ile Ile Arg Val
35 40 45

Asp Pro Thr Cys Pro Leu Ser Ser Asn Pro Gly Thr Gln Val Tyr Glu

34

H

29.06.99

50

55

60

Asp Tyr Asn Cys Thr Leu Asn Gln Thr Asn Ile Glu Asn Asn Asn Asn
65 70 75 80

Lys Phe Tyr Ile Ile Gln Leu Leu Gln Asp Ser Asn Arg Phe Phe Thr
85 90 95

Cys Trp Asn Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Ile
100 105 110

Asn His Phe Thr Arg Leu Glu Asp Ala Lys Lys Asp Phe Glu Lys Lys
115 120 125

Phe Arg Glu Lys Thr Lys Asn Asn Trp Ala Glu Arg Asp His Phe Val
130 135 140

Ser His Pro Gly Lys Tyr Thr Leu Ile Glu Val Gln Ala Glu Asp Glu
145 150 155 160

Ala Gln Glu Ala Val Val Lys Val Asp Arg Gly Pro Val Arg Thr Val
165 170 175

Thr Lys Arg Val Gln Pro Cys Ser Leu Asp Pro Ala Thr Gln Lys Leu
180 185 190

Ile Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Thr Met Ala Leu
195 200 205

Met Asp Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Ser Lys Gln
210 215 220

Gln Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala Leu
225 230 235 240

Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser Ser His
245 250 255

Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro Pro Pro
260 265 270

Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu Leu Val
275 280 285

Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser Glu Gln
290 295 300

35

H

25.08.99

Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln
305 310 315 320

Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro Glu Tyr
325 330 335

Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His Arg Cys
340 345 350

Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu Glu Asp
355 360 365

Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu Trp His
370 375 380

Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg
385 390 395 400

Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala
405 410 415

Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys Cys Gly
420 425 430

Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Arg
435 440 445

Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro Pro Pro
450 455 460

Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp Pro Thr
465 470 475 480

Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val Pro Gln
485 490 495

Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser Gln
500 505 510

Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr Leu
515 520 525

Leu Glu Val His Leu *
530

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:

M 29.06.99

- (A) LÄNGE: 2265 Basenpaare
(B) ART: Nucleotid
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:
(F) GEWEBETYP: Uterus

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) LAGE: 221..1843

(D) SONSTIGE ANGABEN: /product= "Poly ADP Ribose
Polymerase"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

TGGGACTGGT CGCCTGACTC GGCCTGCCCC AGCCTCTGCT TCACCCCACT GGTGGCCAAA

TAGCCGATGT CTAATCCCCC ACACAAGCTC ATCCCCGGCC TCTGGGATTG TTGGGAATTC

TCTCCCTAAT TCACGCCTGA GGCTCATGGA GAGTTGCTAG ACCTGGGACT GCCCTGGGAG

GCGCACACAA CCAGGCCGGG TGGCAGCCAG GACCTCTCCC ATG TCC CTG CTT TTC
Met Ser Leu Leu Phe
535

TTG GCC ATG GCT CCA AAG CCG AAG CCC TGG GTA CAG ACT GAG GGC CCT
Leu Ala Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro
540 545 550 555

GAG AAG AAG AAG GGC CGG CAG GCA GGA AGG GAG GAG GAC CCC TTC CGC
Glu Lys Lys Lys Gly Arg Gln Ala Gly Arg Glu Glu Asp Pro Phe Arg
560 565 570

TCC ACC GCT GAG GCC CTC AAG GCC ATA CCC GCA GAG AAG CGC ATA ATC
Ser Thr Ala Glu Ala Leu Lys Ala Ile Pro Ala Glu Lys Arg Ile Ile
575 580 585

CGC GTG GAT CCA ACA TGT CCA CTC AGC AGC AAC CCC GGG ACC CAG GTG
Arg Val Asp Pro Thr Cys Pro Leu Ser Ser Asn Pro Gly Thr Gln Val

33

H

29.05.99

CAG GGC CAG CCT GTG CCC TGC CCA GAG TTC AGC AGC TCC ACA TTC TCC	1774
Gln Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser	
1070 1075 1080	
CAG AGC GAG TAC CTC ATC TAC CAG GAG AGC CAG TGT CGC CTG CGC TAC	1822
Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr	
1085 1090 1095	
CTG CTG GAG GTC CAC CTC TGA GTGCCCCGCC TGTCCCCCGG GGTCTGCAA	1873
Leu Leu Glu Val His Leu *	
1100 1105	
GGCTGGACTG TGATCTTCAA TCATCCTGCC CATCTCTGGT ACCCCTATAT CACTCCTTTT	1933
TTTCAAGAAT ACAATACGTT GTTGTTAACT ATAGTCACCA TGCTGTACAA GATCCCTGAA	1993
CTTATGCCTC CTAAGTAAA TTTTGTATTC TTTGACACAT CTGCCCAGTC CCTCTCCTCC	2053
CAGCCCATGG TAACCAGCAT TTGACTCTTT ACTTGTATAA GGGCAGCTTT TATAGGTTCC	2113
ACATGTAAGT GAGATCATGC AGTGTGTTGTC TTTCTGTGCC TGGCTTATTT CACTCAGCAT	2173
AATGTGCACC GGGTTCACCC ATGTTTTTCAT AAATGACAAG ATTTCTCCTT TAAAAA	2233
AAAAA	2265

(2) ANGABEN ZU SEQ ID NO: 4:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 534 Aminosäuren

(B) ART: Aminosäure

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

Met	Ala	Pro	Lys	Pro	Lys	Pro	Trp	Val	Gln	Thr	Glu	Gly	Pro	Glu	Lys
1				5					10					15	
Lys	Lys	Gly	Arg	Gln	Ala	Gly	Arg	Glu	Glu	Asp	Pro	Phe	Arg	Ser	Thr
			20					25					30		
Ala	Glu	Ala	Leu	Lys	Ala	Ile	Pro	Ala	Glu	Lys	Arg	Ile	Ile	Arg	Val
			35				40						45		
Asp	Pro	Thr	Cys	Pro	Leu	Ser	Ser	Asn	Pro	Gly	Thr	Gln	Val	Tyr	Glu

H

25.06.99

34

50

55

60

Asp Tyr Asn Cys Thr Leu Asn Gln Thr Asn Ile Glu Asn Asn Asn Asn
65 70 75 80

Lys Phe Tyr Ile Ile Gln Leu Leu Gln Asp Ser Asn Arg Phe Phe Thr
85 90 95

Cys Trp Asn Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser Lys Ile
100 105 110

Asn His Phe Thr Arg Leu Glu Asp Ala Lys Lys Asp Phe Glu Lys Lys
115 120 125

Phe Arg Glu Lys Thr Lys Asn Asn Trp Ala Glu Arg Asp His Phe Val
130 135 140

Ser His Pro Gly Lys Tyr Thr Leu Ile Glu Val Gln Ala Glu Asp Glu
145 150 155 160

Ala Gln Glu Ala Val Val Lys Val Asp Arg Gly Pro Val Arg Thr Val
165 170 175

Thr Lys Arg Val Gln Pro Cys Ser Leu Asp Pro Ala Thr Gln Lys Leu
180 185 190

Ile Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Thr Met Ala Leu
195 200 205

Met Asp Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Ser Lys Gln
210 215 220

Gln Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu Ala Leu
225 230 235 240

Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser Ser His
245 250 255

Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro Pro Pro
260 265 270

Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu Leu Val
275 280 285

Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser Glu Gln
290 295 300

35

M

25.05.99

Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp Tyr Gln
305 310 315 320

Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro Glu Tyr
325 330 335

Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His Arg Cys
340 345 350

Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu Glu Asp
355 360 365

Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu Trp His
370 375 380

Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly Leu Arg
385 390 395 400

Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr Phe Ala
405 410 415

Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys Cys Gly
420 425 430

Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu Gly Arg
435 440 445

Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro Pro Pro
450 455 460

Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp Pro Thr
465 470 475 480

Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val Pro Gln
485 490 495

Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe Ser Gln
500 505 510

Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg Tyr Leu
515 520 525

Leu Glu Val His Leu *
530

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:

H

005099

- (A) LÄNGE: 2265 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

(F) GEWEBETYP: Uterus

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) LAGE: 221..1843

(D) SONSTIGE ANGABEN: /product= "Poly ADP Ribose
Polymerase"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

TGGGACTGGT CGCCTGACTC GGCCTGCCCC AGCCTCTGCT TCACCCCACT GGTGGCCAAA	60
TAGCCGATGT CTAATCCCC ACACAAGCTC ATCCCCGGCC TCTGGGATTG TTGGGAATTC	120
TCTCCCTAAT TCACGCCTGA GGCTCATGGA GAGTTGCTAG ACCTGGGACT GCCCTGGGAG	180
GCGCACACAA CCAGGCCGGG TGGCAGCCAG GACCTCTCCC ATG TCC CTG CTT TTC	235
Met Ser Leu Leu Phe	
535	
TTG GCC ATG GCT CCA AAG CCG AAG CCC TGG GTA CAG ACT GAG GGC CCT	283
Leu Ala Met Ala Pro Lys Pro Lys Pro Trp Val Gln Thr Glu Gly Pro	
540 545 550 555	
GAG AAG AAG AAG GGC CGG CAG GCA GGA AGG GAG GAG GAC CCC TTC CGC	331
Glu Lys Lys Lys Gly Arg Gln Ala Gly Arg Glu Glu Asp Pro Phe Arg	
560 565 570	
TCC ACC GCT GAG GCC CTC AAG GCC ATA CCC GCA GAG AAG CGC ATA ATC	379
Ser Thr Ala Glu Ala Leu Lys Ala Ile Pro Ala Glu Lys Arg Ile Ile	
575 580 585	
CGC GTG GAT CCA ACA TGT CCA CTC AGC AGC AAC CCC GGG ACC CAG GTG	427
Arg Val Asp Pro Thr Cys Pro Leu Ser Ser Asn Pro Gly Thr Gln Val	

M

29.08.99

37

590

595

600

TAT GAG GAC TAC AAC TGC ACC CTG AAC CAG ACC AAC ATC GAG AAC AAC	475
Tyr Glu Asp Tyr Asn Cys Thr Leu Asn Gln Thr Asn Ile Glu Asn Asn	
605 610 615	
AAC AAC AAG TTC TAC ATC ATC CAG CTG CTC CAA GAC AGC AAC CGC TTC	523
Asn Asn Lys Phe Tyr Ile Ile Gln Leu Leu Gln Asp Ser Asn Arg Phe	
620 625 630 635	
TTC ACC TGC TGG AAC CGC TGG GGC CGT GTG GGA GAG GTC GGC CAG TCA	571
Phe Thr Cys Trp Asn Arg Trp Gly Arg Val Gly Glu Val Gly Gln Ser	
640 645 650	
AAG ATC AAC CAC TTC ACA AGG CTA GAA GAT GCA AAG AAG GAC TTT GAG	619
Lys Ile Asn His Phe Thr Arg Leu Glu Asp Ala Lys Lys Asp Phe Glu	
655 660 665	
AAG AAA TTT CGG GAA AAG ACC AAG AAC AAC TGG GCA GAG CGG GAC CAC	667
Lys Lys Phe Arg Glu Lys Thr Lys Asn Asn Trp Ala Glu Arg Asp His	
670 675 680	
TTT GTG TCT CAC CCG GGC AAG TAC ACA CTT ATC GAA GTA CAG GCA GAG	715
Phe Val Ser His Pro Gly Lys Tyr Thr Leu Ile Glu Val Gln Ala Glu	
685 690 695	
GAT GAG GCC CAG GAA GCT GTG GTG AAG GTG GAC AGA GGC CCA GTG AGG	763
Asp Glu Ala Gln Glu Ala Val Val Lys Val Asp Arg Gly Pro Val Arg	
700 705 710 715	
ACT GTG ACT AAG CGG GTG CAG CCC TGC TCC CTG GAC CCA GCC ACG CAG	811
Thr Val Thr Lys Arg Val Gln Pro Cys Ser Leu Asp Pro Ala Thr Gln	
720 725 730	
AAG CTC ATC ACT AAC ATC TTC AGC AAG GAG ATG TTC AAG AAC ACC ATG	859
Lys Leu Ile Thr Asn Ile Phe Ser Lys Glu Met Phe Lys Asn Thr Met	
735 740 745	
GCC CTC ATG GAC CTG GAT GTG AAG AAG ATG CCC CTG GGA AAG CTG AGC	907
Ala Leu Met Asp Leu Asp Val Lys Lys Met Pro Leu Gly Lys Leu Ser	
750 755 760	
AAG CAA CAG ATT GCA CGG GGT TTC GAG GCC TTG GAG GCG CTG GAG GAG	955
Lys Gln Gln Ile Ala Arg Gly Phe Glu Ala Leu Glu Ala Leu Glu Glu	
765 770 775	
GCC CTG AAA GGC CCC ACG GAT GGT GGC CAA AGC CTG GAG GAG CTG TCC	1003

38

Ala Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser Leu Glu Glu Leu Ser
780 785 790 795

TCA CAC TTT TAC ACC GTC ATC CCG CAC AAC TTC GGC CAC AGC CAG CCC 1051
Ser His Phe Tyr Thr Val Ile Pro His Asn Phe Gly His Ser Gln Pro
800 805 810

CCG CCC ATC AAT TCC CCT GAG CTT CTG CAG GCC AAG AAG GAC ATG CTG 1099
Pro Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala Lys Lys Asp Met Leu
815 820 825

CTG GTG CTG GCG GAC ATC GAG CTG GCC CAG GCC CTG CAG GCA GTC TCT 1147
Leu Val Leu Ala Asp Ile Glu Leu Ala Gln Ala Leu Gln Ala Val Ser
830 835 840

GAG CAG GAG AAG ACG GTG GAG GAG GTG CCA CAC CCC CTG GAC CGA GAC 1195
Glu Gln Glu Lys Thr Val Glu Glu Val Pro His Pro Leu Asp Arg Asp
845 850 855

TAC CAG CTT CTC AAG TGC CAG CTG CAG CTG CTA GAC TCT GGA GCA CCT 1243
Tyr Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu Asp Ser Gly Ala Pro
860 865 870 875

GAG TAC AAG GTG ATA CAG ACC TAC TTA GAA CAG ACT GGC AGC AAC CAC 1291
Glu Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln Thr Gly Ser Asn His
880 885 890

AGG TGC CCT ACA CTT CAA CAC ATC TGG AAA GTA AAC CAA GAA GGG GAG 1339
Arg Cys Pro Thr Leu Gln His Ile Trp Lys Val Asn Gln Glu Gly Glu
895 900 905

GAA GAC AGA TTC CAG GCC CAC TCC AAA CTG GGT AAT CGG AAG CTG CTG 1387
Glu Asp Arg Phe Gln Ala His Ser Lys Leu Gly Asn Arg Lys Leu Leu
910 915 920

TGG CAT GGC ACC AAC ATG GCC GTG GTG GCC GCC ATC CTC ACT AGT GGG 1435
Trp His Gly Thr Asn Met Ala Val Val Ala Ala Ile Leu Thr Ser Gly
925 930 935

CTC CGC ATC ATG CCA CAT TCT GGT GGG CGT GTT GGC AAG GGC ATC TAC 1483
Leu Arg Ile Met Pro His Ser Gly Gly Arg Val Gly Lys Gly Ile Tyr
940 945 950 955

TTT GCC TCA GAG AAC AGC AAG TCA GCT GGA TAT GTT ATT GGC ATG AAG 1531
Phe Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr Val Ile Gly Met Lys
960 965 970

H

23.08.99

39

TGT GGG GCC CAC CAT GTC GGC TAC ATG TTC CTG GGT GAG GTG GCC CTG	1579
Cys Gly Ala His His Val Gly Tyr Met Phe Leu Gly Glu Val Ala Leu	
975 980 985	
GGC AGA GAG CAC CAT ATC AAC ACG GAC AAC CCC AGC TTG AAG AGC CCA	1627
Gly Arg Glu His His Ile Asn Thr Asp Asn Pro Ser Leu Lys Ser Pro	
990 995 1000	
CCT CCT GGC TTC GAC AGT GTC ATT GCC CGA GGC CAC ACC GAG CCT GAT	1675
Pro Pro Gly Phe Asp Ser Val Ile Ala Arg Gly His Thr Glu Pro Asp	
1005 1010 1015	
CCG ACC CAG GAC ACT GAG TTG GAG CTG GAT GGC CAG CAA GTG GTG GTG	1723
Pro Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly Gln Gln Val Val Val	
1020 1025 1030 1035	
CCC CAG GGC CAG CCT GTG CCC TGC CCA GAG TTC AGC AGC TCC ACA TTC	1771
Pro Gln Gly Gln Pro Val Pro Cys Pro Glu Phe Ser Ser Ser Thr Phe	
1040 1045 1050	
TCC CAG AGC GAG TAC CTC ATC TAC CAG GAG AGC CAG TGT CGC CTG CGC	1819
Ser Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser Gln Cys Arg Leu Arg	
1055 1060 1065	
TAC CTG CTG GAG GTC CAC CTC TGA GTGCCCCGCC TGTCCCCCGG GGTCTGCAA	1873
Tyr Leu Leu Glu Val His Leu *	
1070 1075	
GGCTGGACTG TGATCTTCAA TCATCCTGCC CATCTCTGGT ACCCCTATAT CACTCCTTTT	1933
TTTCAAGAAT ACAATACGTT GTTGTTAACT ATAGTCACCA TGCTGTACAA GATCCCTGAA	1993
CTTATGCCTC CTAAGTAAAA TTTTGTATTC TTTGACACAT CTGCCCAGTC CCTCTCCTCC	2053
CAGCCCATGG TAACCAGCAT TTGACTCTTT ACTTGTATAA GGGCAGCTTT TATAGGTTCC	2113
ACATGTAAGT GAGATCATGC AGTGTTTGTC TTTCTGTGCC TGGCTTATTT CACTCAGCAT	2173
AATGTGCACC GGGTTCACCC ATGTTTTTCAT AAATGACAAG ATTTCCCTCCT TAAAAA	2233
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA	2265

(2) ANGABEN ZU SEQ ID NO: 6:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 541 Aminosäuren

(B) ART: Aminosäure

H

1908.99

40

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

Met Ser Leu Leu Phe Leu Ala Met Ala Pro Lys Pro Lys Pro Trp Val
1 5 10 15

Gln Thr Glu Gly Pro Glu Lys Lys Lys Gly Arg Gln Ala Gly Arg Glu
20 25 30

Glu Asp Pro Phe Arg Ser Thr Ala Glu Ala Leu Lys Ala Ile Pro Ala
35 40 45

Glu Lys Arg Ile Ile Arg Val Asp Pro Thr Cys Pro Leu Ser Ser Asn
50 55 60

Pro Gly Thr Gln Val Tyr Glu Asp Tyr Asn Cys Thr Leu Asn Gln Thr
65 70 75 80

Asn Ile Glu Asn Asn Asn Asn Lys Phe Tyr Ile Ile Gln Leu Leu Gln
85 90 95

Asp Ser Asn Arg Phe Phe Thr Cys Trp Asn Arg Trp Gly Arg Val Gly
100 105 110

Glu Val Gly Gln Ser Lys Ile Asn His Phe Thr Arg Leu Glu Asp Ala
115 120 125

Lys Lys Asp Phe Glu Lys Lys Phe Arg Glu Lys Thr Lys Asn Asn Trp
130 135 140

Ala Glu Arg Asp His Phe Val Ser His Pro Gly Lys Tyr Thr Leu Ile
145 150 155 160

Glu Val Gln Ala Glu Asp Glu Ala Gln Glu Ala Val Val Lys Val Asp
165 170 175

Arg Gly Pro Val Arg Thr Val Thr Lys Arg Val Gln Pro Cys Ser Leu
180 185 190

Asp Pro Ala Thr Gln Lys Leu Ile Thr Asn Ile Phe Ser Lys Glu Met
195 200 205

Phe Lys Asn Thr Met Ala Leu Met Asp Leu Asp Val Lys Lys Met Pro
210 215 220

H

23.11.99

41

Leu Gly Lys Leu Ser Lys Gln Gln Ile Ala Arg Gly Phe Glu Ala Leu
 225 230 235 240

Glu Ala Leu Glu Glu Ala Leu Lys Gly Pro Thr Asp Gly Gly Gln Ser
 245 250 255

Leu Glu Glu Leu Ser Ser His Phe Tyr Thr Val Ile Pro His Asn Phe
 260 265 270

Gly His Ser Gln Pro Pro Pro Ile Asn Ser Pro Glu Leu Leu Gln Ala
 275 280 285

Lys Lys Asp Met Leu Leu Val Leu Ala Asp Ile Glu Leu Ala Gln Ala
 290 295 300

Leu Gln Ala Val Ser Glu Gln Glu Lys Thr Val Glu Glu Val Pro His
 305 310 315 320

Pro Leu Asp Arg Asp Tyr Gln Leu Leu Lys Cys Gln Leu Gln Leu Leu
 325 330 335

Asp Ser Gly Ala Pro Glu Tyr Lys Val Ile Gln Thr Tyr Leu Glu Gln
 340 345 350

Thr Gly Ser Asn His Arg Cys Pro Thr Leu Gln His Ile Trp Lys Val
 355 360 365

Asn Gln Glu Gly Glu Glu Asp Arg Phe Gln Ala His Ser Lys Leu Gly
 370 375 380

Asn Arg Lys Leu Leu Trp His Gly Thr Asn Met Ala Val Val Ala Ala
 385 390 395 400

Ile Leu Thr Ser Gly Leu Arg Ile Met Pro His Ser Gly Gly Arg Val
 405 410 415

Gly Lys Gly Ile Tyr Phe Ala Ser Glu Asn Ser Lys Ser Ala Gly Tyr
 420 425 430

Val Ile Gly Met Lys Cys Gly Ala His His Val Gly Tyr Met Phe Leu
 435 440 445

Gly Glu Val Ala Leu Gly Arg Glu His His Ile Asn Thr Asp Asn Pro
 450 455 460

Ser Leu Lys Ser Pro Pro Pro Gly Phe Asp Ser Val Ile Ala Arg Gly

BASF Aktiengesellschaft

980301

O.Z. 0050/49100

H

25.09.99

42

465

470

475

480

His Thr Glu Pro Asp Pro Thr Gln Asp Thr Glu Leu Glu Leu Asp Gly
485 490 495

Gln Gln Val Val Val Pro Gln Gly Gln Pro Val Pro Cys Pro Glu Phe
500 505 510

Ser Ser Ser Thr Phe Ser Gln Ser Glu Tyr Leu Ile Tyr Gln Glu Ser
515 520 525

Gln Cys Arg Leu Arg Tyr Leu Leu Glu Val His Leu *
530 535 540

M

28.08.99

Patentansprüche

1. Poly(ADP-ribose)polymerase (PARP)-Homologes, gekennzeichnet
5 durch eine Aminosäuresequenz, welche
a) eine funktionale NAD⁺-Bindungsdomäne und
b) kein Zink-Finger-Sequenzmotiv der allgemeinen Formel



10

aufweist, worin

m für einen ganzzahligen Wert von 28 oder 30 steht und
die Reste X unabhängig voneinander für eine beliebige
Aminosäure stehen;

15

und die funktionalen Äquivalente davon.

2. PARP-Homologes nach Anspruch 1, dadurch gekennzeichnet, daß
die funktionale NAD⁺-Bindungsdomäne folgendes allgemeines Se-
quenzmotiv umfaßt:

20



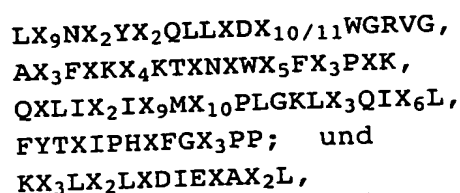
worin

n für einen ganzzahligen Wert von 1 bis 5 steht und die Reste
X unabhängig voneinander für eine beliebige Aminosäure ste-
hen.

25

3. PARP-Homologes nach einem der vorherigen Ansprüche, umfassend
wenigstens ein weiteres der folgenden Teilsequenz-Motive:

30



35

worin die Reste X unabhängig voneinander für eine beliebige
Aminosäure stehen.

- 40 4. Humanes PARP-Homologes nach einem der vorherigen Ansprüche,
gekennzeichnet durch die Aminosäuresequenz gemäß SEQ ID NO:2
(humanPARP2) oder SEQ ID NO:4 oder 6 (humanPARP3 Typ 1 bzw.
2), und die funktionalen Äquivalente davon.

- 45 5. Bindungspartner für PARP-Homologe nach einem der vorherigen
Ansprüche, ausgewählt unter
a) Antikörpern und Fragmenten davon,

c1) an das immobilisierte PARP-Homologe gebundene
teile des Analyten, gegebenenfalls nach einer
Inkubationsphase, bestimmt;

5 oder

a2) einen Analyten, welcher wenigstens einen Bindungs-
partner für das PARP-Homologe enthält
ger immobilisiert;

10 b2) den immobilisierten Analyten mit wenigsten
mologen in Kontakt bringt, für welches man
Bindungspartner sucht; und

c3) den immobilisierten Analyten, gegebenenfalls
Inkubationsphase, auf die Bindung des PARP
tersucht.

15

14. Verfahren zur qualitativen oder quantitativen
PARP-Homologe nach einem der Ansprüche 1 bis 4
Nukleinsäuren, gekennzeichnet durch

20

a) Inkubation einer biologischen Probe mit einer
Menge einer exogenen Nukleinsäure nach einem der
Ansprüche 6 und 7, Hybridisierung unter stringen-
ten, Bestimmung der hybridisierenden Nukleinsäure
gegebenenfalls Vergleich mit einem Standard

25

b) Inkubation einer biologischen Probe mit einem
Oligonucleotid-Primern mit Spezifität für eine
kodierende Nukleinsäure, Amplifizierung der
Nukleinsäure, Bestimmung des Amplifikationsprodukts
gegebenenfalls Vergleich mit einem Standard.

30

15. Verfahren zur qualitativen oder quantitativen
PARP-Homologen nach einem der Ansprüche 1 bis 4
gekennzeichnet durch

35

a) Inkubation einer biologischen Probe mit einem
PARP-Homologes spezifischen Bindungspartner
b) Nachweis des Bindungspartner/PARP-Komplexes
gegebenenfalls
c) Vergleich des Ergebnisses mit einem Standard

40

16. Verfahren nach Anspruch 15, dadurch gekennzeichnet,
dass ein Bindungspartner ein Antikörper oder ein Nukleinsäure
davon ist, der gegebenenfalls eine detektierbare
Träger trägt.

45

M/39113

Majority
490
480
470
460
450
440
430
humanPARP1
humanPARP2
humanPARP3

421
68
17
Majority

Majority

560
550
540
530
520
510
500
humanPARP1
humanPARP2
humanPARP3

491
68
17
Majority

Majority

630
620
610
600
590
580
570
humanPARP1
humanPARP2
humanPARP3

560
109
80
Majority

Majority

700
690
680
670
660
650
640
humanPARP1
humanPARP2
humanPARP3

630
179
148
Majority

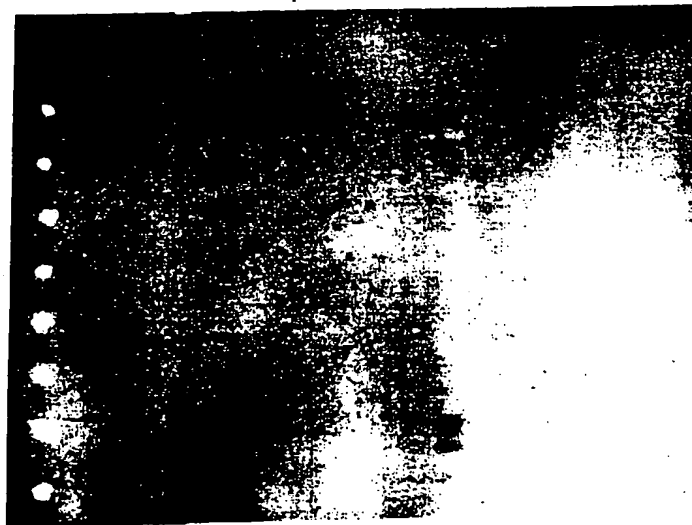
Majority

770
760
750
740
730
720
710
humanPARP1
humanPARP2

691
247

Fig. 2

1 2 3 4 5 6 7 8



— 9,5kb —
— 7,5kb —
— 4,4kb —
— 2,4kb —
— 1,4kb —

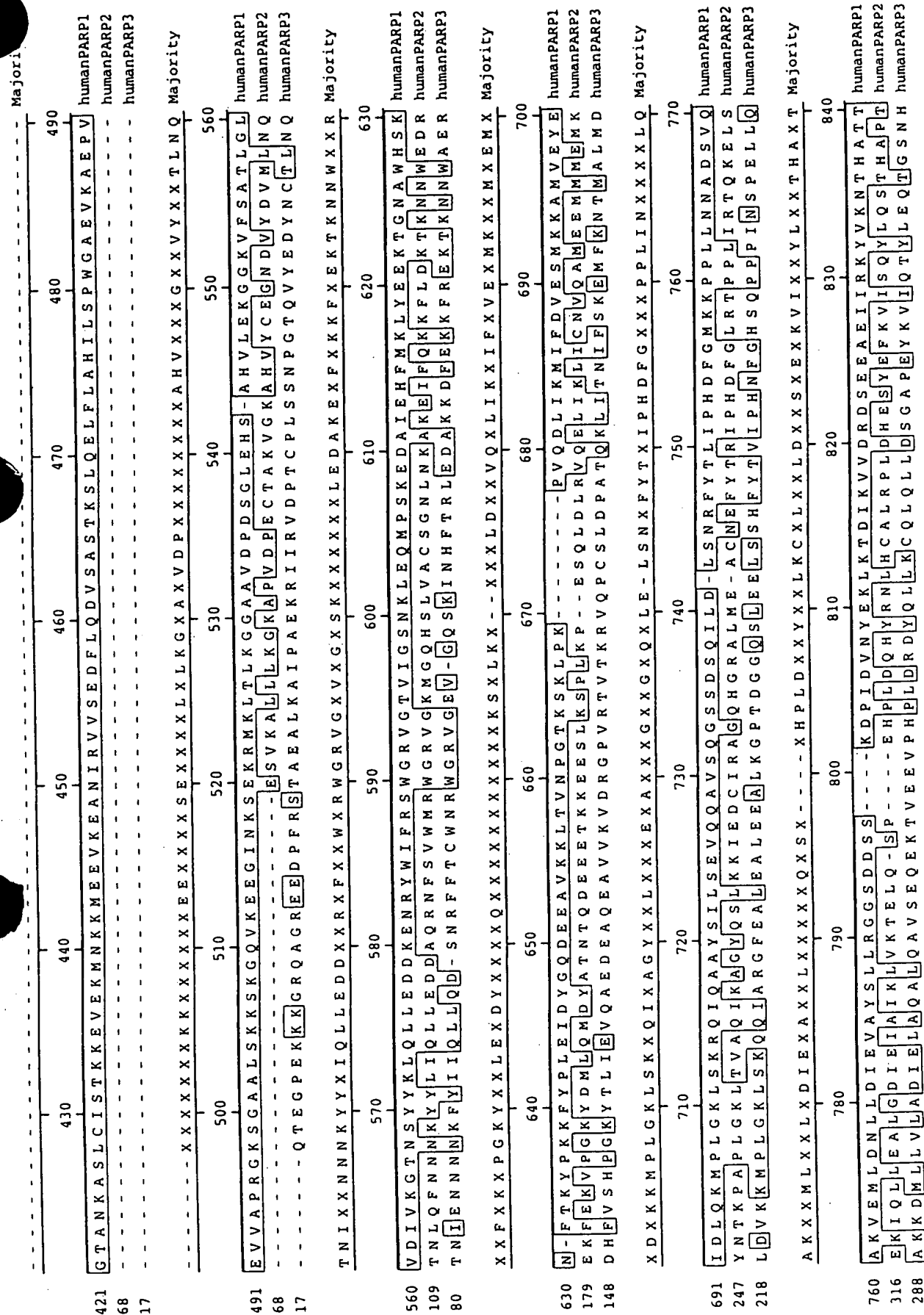
8 7 6 5 4 3 2 1



3 2

	MAXXX	10	20	30	40	50	60	70	Majority	
1	MAESSDKLYRVEYAKSERASCCKKCESI	PKDSLRMAIMVQSPMFDGKVPHWYHFS	CFWKVGHSIRHPDVE						humanPARP1	
1	MAARR								humanPARP2	
1	MSL								humanPARP3	
		80	90	100	110	120	130	140	Majority	
71	VDGFSELRWDDQ	KKTAAGGVTKGQDGIGSKAEKTLGDFAAEYAKSNRSTCKGCMEKIEKGQVRLS							humanPARP1	
6									humanPARP2	
4									humanPARP3	
	XXXXXP	XXXXXXXXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	Majority	
		150	160	170	180	190	200	210	Majority	
141	KKMVDPEK	PQLGMI	DRWYHPGCFVKNR	EELGFRPEYSASQLKGFSL	LATEDK	EALKKQLPGVKSE	GKRKG		humanPARP1	
25	NGNTAPE	DSSPAKKTR	RCQR			QESKKMPVAG	CKANKDR	TE	humanPARP2	
4									humanPARP3	
	DKXD								Majority	
		220	230	240	250	260	270	280	Majority	
211	DKVDGVDE	VAKKSKKEKD	KSKLEKALKAQNDLI	WNIKDELKKVCSTNDL	KELLIFNKQ	QVP	SGESA	IL	humanPARP1	
64	DKQD								humanPARP2	
4									humanPARP3	
									Majority	
		290	300	310	320	330	340	350	Majority	
281	DRVADGMV	FGALLPCE	ECGQLVFKSD	AYCTGDTVATKCMVK	TQTPNRK	EWVTPKE	FREISY	LKKLV	humanPARP1	
68									humanPARP2	
4									humanPARP3	
									Majority	
		360	370	380	390	400	410	420	Majority	
351	KKQDRIF	PPETSA	SVAATPPP	STASAPAAVNSSASADK	PLSNMKILT	LKKLSRN	KKDEV	KAMIEK	LGGKLT	humanPARP1
68										humanPARP2
17										humanPARP3

Fig. 1 (1)

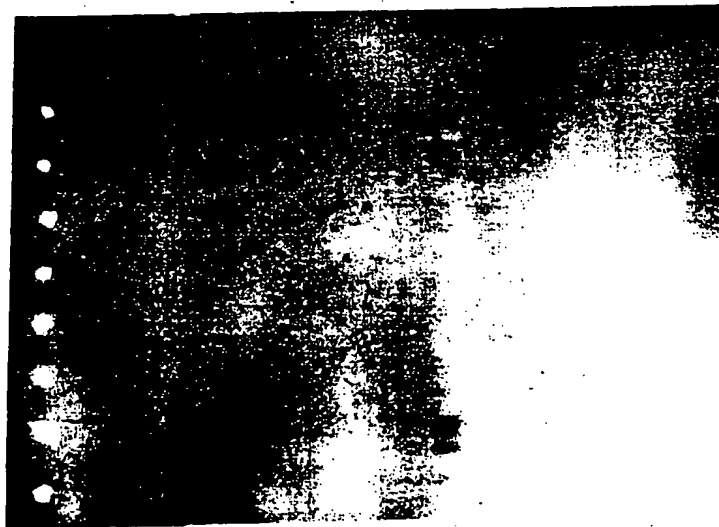


H X Y X X T L X D I F K V E X E G E X X X X X L H N R X L L W H G S R M X X N X A G X G L R I A P P E A P X T G Y M F G K G I Major
 850 860 870 880 890 900 910
 826 H N A Y D L E V I D I F K I E R E G E C Q R Y K P F K Q L H N R R L L W H G S R T T N F A G I L S Q G L R I A P P E A P V T G Y M F G K G I humanPARP1
 381 H S D Y T M T L L D L F E V E K D G E K E A F R - - E D L H N R M L L W H G S R M S N W V G I L S H G L R I A P P E A P I T G Y M F G K G I humanPARP2
 358 R C P - - T L Q H I W K I V N Q E G E D R F Q A H S K L G N R K L L W H G T N M A V V A A I L T S G L R I M P H - - - S G G R V G K G I humanPARP3
 Y F A D M X S K S A N Y C X X S X - - X X X X G L X L L G E V A L G X X X E L X X A N P X X X K X L P X G K H S V K G L G K T X X P D P X X X - Majority
 920 930 940 950 960 970 980
 896 Y F A D M V S K S A N Y C H T S Q - - G D P I G L I L L G E V A L G N M Y E L K K H A S H I S K - L P K G K H S V K G L G K T T P D P S A N - humanPARP1
 449 Y F A D M S S K S A N Y C F A S R - - L K N T G L L L S E V A L G Q C N E L L E A N P K A E G L L Q G K H S T K G L G K M A P S S A H F - humanPARP2
 421 Y F A S E N S K S A G Y V I G M K C G A H H V G Y M F L G E V A L G R E H H I N T D N P S L K S P P G F D S V I A R G H T E P D P T Q D T humanPARP3
 - X X L D G X X V P L G X X X G X X X X X T L X Y N E Y I V Y X X X Q V R L R Y L L K V X F N F X X X L W - Majority
 990 1000 1010 1020 1030
 962 - I S L D G V D V P L G T G I S S G V - - - N D T S L L Y N E Y I V Y D I A Q V N L K Y L L K L K F N F K T S L W - humanPARP1
 516 - V T L N G S T V P L G P A S D T G I L N P D G Y T L N Y N E Y I V Y N P N Q V R M R Y L L K V Q F N F - L Q L W - humanPARP2
 491 E L E L D G Q Q V V V P Q G Q P V P C P E F S S T F S Q S E Y L I Y Q E S Q C R R Y L L E V H L - humanPARP3

Fig. 1 (3)

Fig. 2

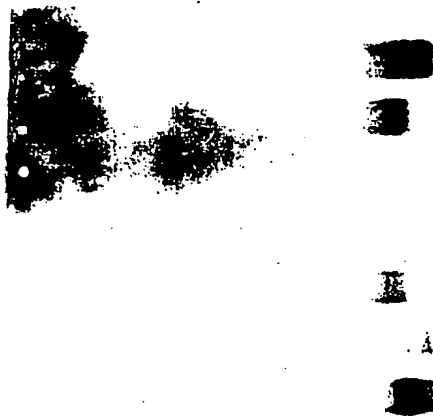
1 2 3 4 5 6 7 8



PARP2

(A)

8 7 6 5 4 3 2 1



PARP3

(B)

— 9,5kb —
— 7,5kb —
— 4,4kb —
— 2,4kb —
— 1,4kb —

1

2 3 4 5 6 7 8